

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:30:47 ; Search time 110.427 Seconds

(without alignments)
2295.451 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 495

Sequence: 1 MKSPVSLAPSDGEGSDRT.....LFAKGEVQNWAI SDHQGRN 495

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : UniProt_03.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	495	100.0	495	1	S175 SHEEP	Qmzdl ovis aries
2	61	12.3	495	1	S175 HUMAN	Qnra2 homo sapien
3	32	6.5	495	1	S175 MOUSE	Q8n82 mus musculu
4	12	2.4	363	2	Q8BWS1	Q8bws1 mus musculu
5	12	2.4	436	1	NPT3 HUMAN	O00624 homo sapien
6	12	2.4	447	2	Q8VCX8	Q8vcx8 mus musculu
7	12	2.4	466	2	Q8GLH1	Q8glh1 homo sapien
8	12	2.4	470	2	Q7YTZ7	Q7ytz7 drosophila
9	12	2.4	559	2	Q9VIG7	Q9vyg7 drosophila
10	11	2.2	582	2	Q9P2U8	Q9p2u8 homo sapien
11	11	2.2	582	2	Q920B7	Q920b7 mus musculu
12	11	2.2	582	2	Q8BLE7	Q8ble7 mus musculu
13	11	2.2	582	2	Q8J1I2	Q8j1i2 rattus norv
14	10	2.0	390	2	Q921B2	Q921b2 mus musculu
15	10	2.0	465	1	NFT1 MOUSE	Q61983 mus musculu
16	10	2.0	465	1	NFT1 RAT	Q62795 rattus norv
17	10	2.0	465	2	Q8K3H3	Q8k3h3 rattus norv
18	10	2.0	465	2	Q91WV5	Q91wv5 mus musculu
19	10	2.0	465	2	Q6AZ46	Q6az46 rattus norv
20	10	2.0	476	2	Q7Q5Y9	Q7q5y9 anopheles g
21	10	2.0	497	2	Q6NR28	Q6nr28 drosophila
22	10	2.0	497	2	Q7KRJ1	Q7krj1 drosophila
23	10	2.0	502	2	Q9VDM0	Q9vdm0 drosophila
24	10	2.0	544	1	YRT3 CAEBL	Q10046 caenorhabdi
25	10	2.0	544	2	Q7Q580	Q7q580 anopheles g
26	10	2.0	573	2	Q9Y932	Q9y932 caenorhabdi
27	9	1.8	69	2	Q91Y99	Q91y99 mus musculu
28	9	1.8	127	2	Q71118	Q71118 trichoplusi
29	9	1.8	395	2	Q65502	Q65502 arabidopsis
30	9	1.8	469	2	Q8CJH8	Q8cjh8 rattus norv
31	9	1.8	475	2	Q6AZ69	Q6az69 rattus norv

32	9	1.8	479	2	Q23514	Q23514 caenorhabdi
33	9	1.8	497	2	Q8MRP7	Q8mrp7 drosophila
34	9	1.8	498	2	Q8CJH9	Q8cjh9 rattus norv
35	8	1.6	119	2	Q6EX40	Q6ex40 pimelia rad
36	8	1.6	126	2	Q9LRG3	Q9lrg3 chlorella v
37	8	1.6	130	2	Q20936	Q20936 simulium vi
38	8	1.6	130	2	Q9T9F5	Q9t9f5 hermatobate
39	8	1.6	131	2	Q9TH57	Q9th57 littorina s
40	8	1.6	133	2	Q9XPE9	Q9xpe9 triatoma pr
41	8	1.6	133	2	Q9XPC8	Q9xpc8 rhodnius pr
42	8	1.6	133	2	Q9XPC9	Q9xpc9 rhodnius ro
43	8	1.6	135	2	Q9ZXN3	Q9zxn3 anopheles f
44	8	1.6	136	2	Q94TM4	Q94tm4 cicindela c
45	8	1.6	136	2	Q9G7H1	Q9g7h1 anopheles n
46	8	1.6	136	2	Q9G7H5	Q9g7h5 uranotaenia
47	8	1.6	136	2	Q9G7H6	Q9g7h6 anopheles c
48	8	1.6	136	2	Q9G7H7	Q9g7h7 anopheles s
49	8	1.6	136	2	Q9G7H8	Q9g7h8 chagasia ba
50	8	1.6	136	2	Q9G7H9	Q9g7h9 anopheles a
51	8	1.6	136	2	Q9G7I0	Q9g7i0 anopheles a
52	8	1.6	136	2	Q9G7I1	Q9g7i1 biranella g
53	8	1.6	136	2	Q9G7I2	Q9g7i2 anopheles b
54	8	1.6	136	2	Q9G7I3	Q9g7i3 anopheles c
55	8	1.6	138	2	Q8UE72	Q8ue72 agrobacteri
56	8	1.6	139	2	Q9XKI8	Q9xki8 nezara viri
57	8	1.6	141	2	Q51581	Q51581 plectonema
58	8	1.6	144	2	Q6VTI1	Q6vti1 lygus genel
59	8	1.6	159	2	Q7YH75	Q7yh75 spinaerodact
60	8	1.6	160	2	Q21745	Q21745 littorina s
61	8	1.6	160	2	Q79238	Q79238 littorina s
62	8	1.6	160	2	Q79239	Q79239 littorina p
63	8	1.6	160	2	Q99830	Q99830 littorina s
64	8	1.6	160	2	Q34932	Q34932 littorina s
65	8	1.6	160	2	Q9T337	Q9t337 littorina s
66	8	1.6	160	2	Q9T356	Q9t356 littorina s
67	8	1.6	160	2	Q9T3N7	Q9t3n7 littorina p
68	8	1.6	160	2	Q9T3Q3	Q9t3q3 littorina p
69	8	1.6	160	2	Q9T5E2	Q9t5e2 littorina p
70	8	1.6	160	2	Q9T5F0	Q9t5f0 littorina p
71	8	1.6	160	2	Q9T5F1	Q9t5f1 littorina s
72	8	1.6	160	2	Q9T5F2	Q9t5f2 littorina s
73	8	1.6	160	2	Q9T5F3	Q9t5f3 littorina s
74	8	1.6	172	2	Q7PWK3	Q7pwk3 anopheles g
75	8	1.6	178	2	Q9T3K8	Q9t3k8 littorina a
76	8	1.6	178	2	Q9T3K9	Q9t3k9 littorina s
77	8	1.6	178	2	Q9T461	Q9t461 littorina f
78	8	1.6	178	2	Q9THC1	Q9thc1 littorina l
79	8	1.6	178	2	Q9THC2	Q9thc2 littorina l
80	8	1.6	178	2	Q9THC3	Q9thc3 littorina o
81	8	1.6	186	2	Q7YAE8	Q7yae8 thalassoma
82	8	1.6	189	2	Q8HBE4	Q8hbe4 littorina l
83	8	1.6	189	2	Q8HNN7	Q8hnn7 littorina l
84	8	1.6	193	2	Q6DTN7	Q6dtn7 drosophila
85	8	1.6	193	2	Q6DTN8	Q6dtn8 drosophila
86	8	1.6	193	2	Q6DTN9	Q6dtn9 drosophila
87	8	1.6	193	2	Q6DTP0	Q6dtp0 drosophila
88	8	1.6	193	2	Q6DTP5	Q6dtp5 drosophila
89	8	1.6	193	2	Q6DTQ6	Q6dtq6 drosophila
90	8	1.6	193	2	Q6DTQ8	Q6dtq8 drosophila
91	8	1.6	205	2	Q6QTE2	Q6qte2 bolitogloss
92	8	1.6	206	2	Q6QTD0	Q6qtd0 bolitogloss
93	8	1.6	208	2	Q6QTD3	Q6qtd3 bolitogloss
94	8	1.6	211	2	Q9ZZU2	Q9zzu2 littorina s
95	8	1.6	221	2	Q8M7R2	Q8m7r2 rhodnius ro
96	8	1.6	221	2	Q8M7R3	Q8m7r3 rhodnius ro
97	8	1.6	221	2	Q8M7R4	Q8m7r4 rhodnius ro
98	8	1.6	221	2	Q8M7R5	Q8m7r5 rhodnius pr
99	8	1.6	221	2	Q8M7R6	Q8m7r6 rhodnius pr
100	8	1.6	222	2	Q34402	Q34402 euproctus a
101	8	1.6	223	2	Q34860	Q34860 littorina a
102	8	1.6	223	2	Q34861	Q34861 littorina a
103	8	1.6	223	2	Q34867	Q34867 littorina b
104	8	1.6	223	2	Q34871	Q34871 littorina c

105	8	1.6	223	2	Q34892	Q34892 littorina h	178	8	1.6	303	2	Q71BT0	Q71b0 blackburnia
106	8	1.6	223	2	Q34899	Q34899 littorina k	179	8	1.6	308	1	LIVH_SALTY	P30295 salmonella
107	8	1.6	223	2	Q34906	Q34906 littorina l	180	8	1.6	311	2	Q71BR7	Q71br7 blackburnia
108	8	1.6	223	2	Q34908	Q34908 littorina m	181	8	1.6	311	2	Q71BT4	Q71bc4 blackburnia
109	8	1.6	223	2	Q34914	Q34914 littorina o	182	8	1.6	316	2	Q71BS5	Q71bs5 blackburnia
110	8	1.6	223	2	Q34924	Q34924 littorina p	183	8	1.6	316	2	Q71BS7	Q71bs7 blackburnia
111	8	1.6	223	2	Q34933	Q34933 littorina s	184	8	1.6	316	2	Q71BS6	Q71bs6 oryza sativ
112	8	1.6	223	2	Q34934	Q34934 littorina s	185	8	1.6	320	2	Q71BS4	Q71bs4 blackburnia
113	8	1.6	223	2	Q7GG43	Q7gg43 littorina s	186	8	1.6	321	2	Q7F919	Q7f919 oryza sativ
114	8	1.6	226	2	Q9B388	Q9b388 blackburnia	187	8	1.6	326	2	Q9B707	Q9b707 eurycea wat
115	8	1.6	228	2	Q81V29	Q81v29 tarentola e	188	8	1.6	327	2	Q71BR1	Q71br1 blackburnia
116	8	1.6	232	2	Q8M376	Q8m376 austrochilus	189	8	1.6	327	2	Q71BR3	Q71br3 blackburnia
117	8	1.6	237	2	Q03761	Q03761 nucella lam	190	8	1.6	327	2	Q71BR5	Q71br5 blackburnia
118	8	1.6	238	2	Q03745	Q03745 nucella can	191	8	1.6	327	2	Q71BR6	Q71br6 blackburnia
119	8	1.6	238	2	Q03750	Q03750 nucella fire	192	8	1.6	327	2	Q71BR8	Q71br8 blackburnia
120	8	1.6	238	2	Q03760	Q03760 nucella lam	193	8	1.6	327	2	Q71BU3	Q71bu3 blackburnia
121	8	1.6	238	2	Q9GCH1	Q9gch1 ceratinosto	194	8	1.6	328	2	Q71BT2	Q71bt2 blackburnia
122	8	1.6	238	2	Q9GCH2	Q9gch2 scathophaga	195	8	1.6	329	2	Q71BS1	Q71bs1 blackburnia
123	8	1.6	238	2	Q9GCH3	Q9gch3 scathophaga	196	8	1.6	330	2	Q71BR8	Q71br8 blackburnia
124	8	1.6	238	2	Q9GCH4	Q9gch4 scathophaga	197	8	1.6	332	2	Q71BS0	Q71bs0 blackburnia
125	8	1.6	238	2	Q9GCH5	Q9gch5 scathophaga	198	8	1.6	332	2	Q9B706	Q9b706 eurycea wat
126	8	1.6	238	2	Q9GCH6	Q9gch6 scathophaga	199	8	1.6	333	2	Q71BT1	Q71bt1 blackburnia
127	8	1.6	238	2	Q9GCH7	Q9gch7 scathophaga	200	8	1.6	338	2	Q71BR0	Q71br0 blackburnia
128	8	1.6	238	2	Q9GCH8	Q9gch8 scathophaga	201	8	1.6	338	2	Q71BR2	Q71br2 blackburnia
129	8	1.6	238	2	Q9GCH9	Q9gch9 scathophaga	202	8	1.6	338	2	Q71BR4	Q71br4 blackburnia
130	8	1.6	238	2	Q9GCI0	Q9gci0 scathophaga	203	8	1.6	338	2	Q71BS6	Q71bs6 blackburnia
131	8	1.6	238	2	Q9GCI1	Q9gci1 scathophaga	204	8	1.6	338	2	Q7XWF4	Q7xwf4 oryza sativ
132	8	1.6	238	2	Q9GCI2	Q9gci2 scathophaga	205	8	1.6	339	2	Q7XUH8	Q7xuh8 oryza sativ
133	8	1.6	238	2	Q9GCI3	Q9gci3 scathophaga	206	8	1.6	341	2	Q9B710	Q9b710 eurycea tro
134	8	1.6	238	2	Q9GCI4	Q9gci4 scathophaga	207	8	1.6	342	2	Q21438	Q21438 caenorhabdi
135	8	1.6	238	2	Q9GCI5	Q9gci5 scathophaga	208	8	1.6	358	2	Q71BU7	Q71bu7 blackburnia
136	8	1.6	238	2	Q9GCI6	Q9gci6 scathophaga	209	8	1.6	359	2	Q71B00	Q71b00 caenorhabdi
137	8	1.6	239	2	Q03746	Q03746 nucella can	210	8	1.6	364	2	Q71BT9	Q71bt9 blackburnia
138	8	1.6	239	2	Q03748	Q03748 nucella ema	211	8	1.6	364	2	Q71BU0	Q71bu0 blackburnia
139	8	1.6	239	2	Q03749	Q03749 nucella ema	212	8	1.6	364	2	Q71BU4	Q71bu4 blackburnia
140	8	1.6	239	2	Q03752	Q03752 nucella fire	213	8	1.6	364	2	Q71BU9	Q71bu9 blackburnia
141	8	1.6	239	2	Q03753	Q03753 nucella lim	214	8	1.6	364	2	Q71BV0	Q71bv0 blackburnia
142	8	1.6	239	2	Q03754	Q03754 nucella lim	215	8	1.6	365	1	GLE2_YEAST	P40066 saccharomyc
143	8	1.6	239	2	Q03757	Q03757 nucella lap	216	8	1.6	369	2	Q72KC7	Q72kc7 thermus the
144	8	1.6	239	2	Q03758	Q03758 nucella lam	217	8	1.6	371	2	Q71BU1	Q71bu1 blackburnia
145	8	1.6	239	2	Q03759	Q03759 nucella lim	218	8	1.6	371	2	Q8HQ06	Q8hq06 thrips imag
146	8	1.6	243	2	Q75YPL	Q75ypl nesiohelix	219	8	1.6	372	2	Q6QRA1	Q6qra1 eurycea tyr
147	8	1.6	248	2	Q8M111	Q8m111 harpagifer	220	8	1.6	372	2	Q6QRB9	Q6qrb9 eurycea tyr
148	8	1.6	255	2	Q9B394	Q9b394 blackburnia	221	8	1.6	372	2	Q6QRE2	Q6qre2 eurycea mul
149	8	1.6	257	2	Q8S624	Q8s624 oryza sativ	222	8	1.6	372	2	Q71BU2	Q71bu2 blackburnia
150	8	1.6	258	2	Q7XLL7	Q7xll7 oryza sativ	223	8	1.6	378	1	CYB_ANOGA	P34844 anopheles g
151	8	1.6	260	2	Q9B380	Q9b380 blackburnia	224	8	1.6	378	1	CYB_ANOQU	P33501 anopheles q
152	8	1.6	260	2	Q9B381	Q9b381 blackburnia	225	8	1.6	378	2	Q8M6E0	Q8m6e0 barbus trim
153	8	1.6	260	2	Q9B389	Q9b389 blackburnia	226	8	1.6	378	2	Q6B3B1	Q6b3b1 pteronarcys
154	8	1.6	260	2	Q9B391	Q9b391 blackburnia	227	8	1.6	380	1	CYB_CEPNE	Q34179 cepaea nemo
155	8	1.6	260	2	Q9B395	Q9b395 blackburnia	228	8	1.6	380	2	Q9B2C3	Q9b2c3 barbus bigo
156	8	1.6	260	2	Q9B398	Q9b398 blackburnia	229	8	1.6	380	2	Q9G0M6	Q9g0m6 osteoglossu
157	8	1.6	260	2	Q9B3A4	Q9b3a4 blackburnia	230	8	1.6	380	2	Q9G0M7	Q9g0m7 osteoglossu
158	8	1.6	260	2	Q9B3B0	Q9b3b0 blackburnia	231	8	1.6	380	2	Q644A1	Q644a1 bolitoglossa
159	8	1.6	260	2	Q9B3B7	Q9b3b7 blackburnia	232	8	1.6	380	2	Q644K5	Q644k5 desmognathu
160	8	1.6	260	2	Q9B3C2	Q9b3c2 blackburnia	233	8	1.6	380	2	Q645C5	Q645c5 aneides fla
161	8	1.6	260	2	Q9B3C7	Q9b3c7 blackburnia	234	8	1.6	381	2	Q35424	Q35424 protopterus
162	8	1.6	260	2	Q9B3D0	Q9b3d0 blackburnia	235	8	1.6	402	2	Q7TQH3	Q7tdh3 mus musculu
163	8	1.6	260	2	Q9B3D3	Q9b3d3 blackburnia	236	8	1.6	454	2	Q8Y0F1	Q8y0f1 ralstonia s
164	8	1.6	260	2	Q9B3D6	Q9b3d6 blackburnia	237	8	1.6	465	1	NPT1_RABIT	Q28722 o renal sod
165	8	1.6	260	2	Q9B3D9	Q9b3d9 blackburnia	238	8	1.6	465	2	Q8VZY0	Q8vzy0 oryza sativ
166	8	1.6	260	2	Q9B3E2	Q9b3e2 notagonum s	239	8	1.6	476	2	Q8ZZR7	Q8zzr7 pyrobaculum
167	8	1.6	269	2	Q58366	Q58366 pyrococcus	240	8	1.6	479	1	DBPA_BACSU	P42305 bacillus su
168	8	1.6	269	2	Q8U2T4	Q8u2t4 pyrococcus	241	8	1.6	494	2	Q83EH5	Q83eh5 coxiella bu
169	8	1.6	269	2	Q8S691	Q8s691 oryza sativ	242	8	1.6	497	2	Q9Y2C5	Q9y2c5 homo sapien
170	8	1.6	269	2	Q7G785	Q7g785 oryza sativ	243	8	1.6	498	2	Q8VCY5	Q8vcy5 mus musculu
171	8	1.6	269	2	Q7XLL9	Q7xll9 oryza sativ	244	8	1.6	502	2	Q7XGE6	Q7xge6 oryza sativ
172	8	1.6	269	2	Q851B3	Q851b3 oryza sativ	245	8	1.6	515	2	Q8JFT2	Q8jft2 brachydanio
173	8	1.6	284	2	Q6U058	Q6u058 oryza sativ	246	8	1.6	535	2	Q7PWK4	Q7pwk4 anopheles g
174	8	1.6	291	2	Q8LHH2	Q8lhh2 oryza sativ	247	8	1.6	560	2	Q6PCD0	Q6pcd0 homo sapien
175	8	1.6	291	2	Q7XKX8	Q7xkx8 oryza sativ	248	8	1.6	560	2	Q9P2U7	Q9p2u7 homo sapien
176	8	1.6	299	2	Q7XK66	Q7xk66 oryza sativ	249	8	1.6	560	2	Q62E34	Q62e34 rattus norv
177	8	1.6	300	2	Q33966	Q33966 brachylophu	250	8	1.6	576	2	Q6INC8	Q6inc8 xenopus lae

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:34:43 ; Search time 24.9661 Seconds
(without alignments)
1907.682 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 495

Sequence: 1 MKSPVSLAPSGEGSDRT.....LFAKGVQNWALSHQGHNR 495

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

PIR 79:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	2.0	465	2 S69915	sodium-phosphate t
2	10	2.0	544	2 T24633	hypothetical prote
3	10	2.0	573	2 T23589	hypothetical prote
4	9	1.8	380	2 T29248	hypothetical prote
5	9	1.8	395	2 T06132	hypothetical prote
6	8	1.6	138	2 E37587	diacylglycerol kin
7	8	1.6	138	2 AG2808	diacylglycerol kin
8	8	1.6	269	1 H71107	guFA protein homol
9	8	1.6	308	2 AG0992	high-affinity bran
10	8	1.6	308	2 JH0668	high-affinity bran
11	8	1.6	359	2 T32287	hypothetical prote
12	8	1.6	365	2 S06010	hypothetical prote
13	8	1.6	378	2 T12009	ubiquinol-cytochro
14	8	1.6	378	2 T09812	ubiquinol-cytochro
15	8	1.6	380	2 T11381	ubiquinol-cytochro
16	8	1.6	381	2 S68140	ubiquinol-cytochro
17	8	1.6	396	2 T23619	hypothetical prote
18	8	1.6	465	2 A56410	sodium/phosphate t
19	8	1.6	479	2 E69613	ATP-dependent RNA
20	8	1.6	560	2 T59302	brain specific Na+
21	8	1.6	837	2 T12514	hypothetical prote
22	8	1.6	1287	2 AF2031	two-component hybr
23	8	1.6	4128	2 JG6306	protein kinase (EC
24	7	1.4	25	2 A60704	serine proteinase
25	7	1.4	33	2 S13863	dihydrolipoamide d
26	7	1.4	44	2 F95124	hypothetical prote
27	7	1.4	55	2 A21124	Sxr (Bkm-homolog)
28	7	1.4	68	2 B64161	yacG protein homol
29	7	1.4	126	2 F82072	holo-(acyl-carrier

30	7	1.4	136	2 G97824	50S ribosomal prot
31	7	1.4	140	2 S09525	mercury transport
32	7	1.4	152	2 PT0433	progesterone 11alp
33	7	1.4	153	2 B82330	conserved hypothet
34	7	1.4	165	2 C97624	petP protein limpo
35	7	1.4	165	2 AD2847	transcription regu
36	7	1.4	169	1 OLBO4	cytochrome-c oxida
37	7	1.4	172	2 AH2313	phosphatidylglycer
38	7	1.4	182	2 T17625	hypothetical prote
39	7	1.4	185	2 JK0021	somatotropin - ski
40	7	1.4	187	2 A61123	somatotropin - Moz
41	7	1.4	187	2 PC1082	somatotropin - Jap
42	7	1.4	187	2 JU0030	somatotropin - yel
43	7	1.4	187	2 PC1087	somatotropin - dus
44	7	1.4	187	2 PC1088	somatotropin - gre
45	7	1.4	191	2 S74679	hypothetical prote
46	7	1.4	192	2 AI3437	colicin v producti
47	7	1.4	197	2 A70750	hypothetical prote
48	7	1.4	203	2 S00747	somatotropin precu
49	7	1.4	204	1 STFI	somatotropin precu
50	7	1.4	204	2 A56642	somatotropin precu
51	7	1.4	204	2 I51289	somatotropin - Mor
52	7	1.4	204	2 JN0484	somatotropin precu
53	7	1.4	204	2 JH0577	somatotropin precu
54	7	1.4	204	2 A56904	somatotropin precu
55	7	1.4	204	2 S30431	somatotropin - Aca
56	7	1.4	204	2 JC4261	somatotropin precu
57	7	1.4	204	2 S01746	somatotropin precu
58	7	1.4	212	2 C70898	hypothetical prote
59	7	1.4	225	2 B83083	conserved hypothet
60	7	1.4	231	2 JC6201	GSI protein homolo
61	7	1.4	236	2 S42069	TEGT protein - rat
62	7	1.4	242	2 AE2314	ATP-binding protei
63	7	1.4	247	2 S72859	hypothetical prote
64	7	1.4	252	2 E90184	conserved hypothet
65	7	1.4	252	2 E82040	general secretion
66	7	1.4	279	2 E64109	dimethylsulfoxide
67	7	1.4	279	2 H70855	hypothetical prote
68	7	1.4	282	2 E95128	conserved hypothet
69	7	1.4	282	2 E97999	conserved hypothet
70	7	1.4	284	1 I76721	hypothetical prote
71	7	1.4	284	1 S60666	probable ATP-bindi
72	7	1.4	284	2 AB0906	conserved hypothet
73	7	1.4	284	2 G85984	hypothetical prote
74	7	1.4	284	2 D91139	hypothetical prote
75	7	1.4	284	2 AB0436	conserved hypothet
76	7	1.4	285	2 S74771	hypothetical prote
77	7	1.4	289	1 G64167	hypothetical prote
78	7	1.4	300	2 AE0762	conserved hypothet
79	7	1.4	301	1 H64438	UDP-N-acetylglucos
80	7	1.4	304	2 H75081	phosphate abc tran
81	7	1.4	305	1 H75285	probable phosphoe
82	7	1.4	307	2 E91206	probable ARAC-type
83	7	1.4	307	2 G86052	probable ARAC-type
84	7	1.4	307	2 A65170	hypothetical prote
85	7	1.4	308	2 G98084	hypothetical prote
86	7	1.4	308	2 G95220	hypothetical prote
87	7	1.4	308	2 S11151	amid protein - Str
88	7	1.4	313	2 B71390	NADH2 dehydrogenas
89	7	1.4	315	2 E64426	phosphate transpor
90	7	1.4	316	2 E70768	hypothetical prote
91	7	1.4	317	2 T11337	NADH2 dehydrogenas
92	7	1.4	318	1 QXBOIM	NADH2 dehydrogenas
93	7	1.4	318	2 A58888	NADH2 dehydrogenas
94	7	1.4	318	2 S47870	NADH2 dehydrogenas
95	7	1.4	318	2 T11428	NADH2 dehydrogenas
96	7	1.4	318	2 T11441	NADH2 dehydrogenas
97	7	1.4	318	2 T11140	NADH2 dehydrogenas
98	7	1.4	318	2 T11493	NADH2 dehydrogenas
99	7	1.4	318	2 T11402	NADH2 dehydrogenas
100	7	1.4	318	2 S41835	NADH2 dehydrogenas
101	7	1.4	318	2 S26151	NADH2 dehydrogenas
102	7	1.4	318	2 A58850	NADH2 dehydrogenas

103	7	1.4	318	2	S41820	NADH2 dehydrogenas	176	7	1.4	559	2	S62503	inorganic phosphat
104	7	1.4	318	2	T45550	NADH2 dehydrogenas	177	7	1.4	563	2	T43650	protein sodium-de
105	7	1.4	318	2	T11363	NADH2 dehydrogenas	178	7	1.4	576	2	H88548	protein ZK512.6 [i
106	7	1.4	318	2	T11857	NADH2 dehydrogenas	179	7	1.4	600	2	E48951	nisin transport pr
107	7	1.4	318	2	T11247	NADH2 dehydrogenas	180	7	1.4	635	2	F70874	probable membrane
108	7	1.4	318	2	T10972	NADH2 dehydrogenas	181	7	1.4	640	2	D83571	conserved hypotnet
109	7	1.4	318	2	T11050	NADH2 dehydrogenas	182	7	1.4	644	2	D90940	hypothetical prote
110	7	1.4	321	2	T13811	NADH2 dehydrogenas	183	7	1.4	644	2	G64938	hypothetical prote
111	7	1.4	321	2	S55004	NADH2 dehydrogenas	184	7	1.4	644	2	H85788	hypothetical prote
112	7	1.4	321	2	AE0775	probable sugar kin	185	7	1.4	644	2	F82145	conserved hypotnet
113	7	1.4	322	2	S68128	NADH2 dehydrogenas	186	7	1.4	644	2	AF0262	conserved hypotnet
114	7	1.4	323	1	QXXLM	NADH2 dehydrogenas	187	7	1.4	648	2	T47988	serine/threonine-p
115	7	1.4	323	2	T11793	NADH2 dehydrogenas	188	7	1.4	663	2	T38155	78 kd glucose regu
116	7	1.4	323	2	A34284	NADH2 dehydrogenas	189	7	1.4	663	2	S20877	dnak-type molecula
117	7	1.4	323	2	A58892	NADH2 dehydrogenas	190	7	1.4	674	2	I55476	growth potentiati
118	7	1.4	323	2	T09857	NADH2 dehydrogenas	191	7	1.4	678	2	B48089	growth arrest-spec
119	7	1.4	323	2	T11820	NADH2 dehydrogenas	192	7	1.4	682	2	T15092	hypothetical prote
120	7	1.4	324	2	S36002	NADH2 dehydrogenas	193	7	1.4	752	2	G02273	Liv-1 protein - hu
121	7	1.4	324	2	S35462	NADH2 dehydrogenas	194	7	1.4	768	2	T27855	hypothetical prote
122	7	1.4	324	2	T09947	NADH2 dehydrogenas	195	7	1.4	778	2	T17679	proline-rich prote
123	7	1.4	324	2	C41608	hypothetical prote	196	7	1.4	781	2	T02272	hypothetical prote
124	7	1.4	324	2	D87075	probable DNA-bindi	197	7	1.4	792	2	S32244	X-Pro dipeptidyl-p
125	7	1.4	326	1	S41120	cholestenone 5beta	198	7	1.4	847	2	G75270	cation-transportin
126	7	1.4	328	2	T36494	probable membrane	199	7	1.4	860	2	G82310	alanyl-tRNA synth
127	7	1.4	330	2	A40602	3',5'-cyclic-nucle	200	7	1.4	877	2	S65057	alpha-glucosidase
128	7	1.4	333	2	E97368	probable enzyme (A	201	7	1.4	909	2	D87434	phosphoenolpyruvat
129	7	1.4	338	2	AD2586	ribosomal RNA smal	202	7	1.4	924	2	S75284	chemotaxis protein
130	7	1.4	343	2	T14245	NADH2 dehydrogenas	203	7	1.4	925	2	E83529	sensor/response re
131	7	1.4	351	2	T19623	hypothetical prote	204	7	1.4	928	2	D72077	polymorphic outer
132	7	1.4	370	2	A80602	probable membrane	205	7	1.4	928	2	H86546	polymorphic outer
133	7	1.4	372	2	S20054	ubiquinol-cytochro	206	7	1.4	949	2	F81591	polymorphic membra
134	7	1.4	372	2	S20056	para-hydroxybenzoa	207	7	1.4	1005	1	P1VXPJ	RNA 1 protein - pe
135	7	1.4	373	2	I37463	acetylserotonin O-	208	7	1.4	1018	2	T15297	hypothetical prote
136	7	1.4	377	2	F95982	probable transport	209	7	1.4	1082	2	T44177	hypothetical prote
137	7	1.4	381	2	T29300	hypothetical prote	210	7	1.4	1082	2	T43990	hypothetical prote
138	7	1.4	384	2	A70805	probable lprn prot	211	7	1.4	1083	2	H86921	probable arabinosy
139	7	1.4	389	2	T10194	hypothetical prote	212	7	1.4	1085	2	F96712	hypothetical prote
140	7	1.4	390	2	T50037	capsular polysacch	213	7	1.4	1086	2	T09325	probable capsid as
141	7	1.4	397	2	S23909	translation elonga	214	7	1.4	1208	2	T05077	hypothetical prote
142	7	1.4	407	2	F70696	hypothetical prote	215	7	1.4	1225	1	B64234	hypothetical prote
143	7	1.4	410	2	S74951	Na+/H+-exchanging	216	7	1.4	1325	1	S73723	probable lipoprote
144	7	1.4	422	2	C82912	phosphoglycerate k	217	7	1.4	1436	2	B81704	conserved hypotnet
145	7	1.4	422	2	AB3635	probable O-antigen	218	7	1.4	1516	2	F83085	conserved hypotnet
146	7	1.4	434	2	G83129	probable MFS trans	219	7	1.4	1538	2	B90924	probable ATP-depen
147	7	1.4	436	2	C69764	4-aminobutyrate am	220	7	1.4	1538	2	F85772	ATP-dependent heli
148	7	1.4	443	2	S77166	hypothetical prote	221	7	1.4	1538	2	G64922	probable ATP-depen
149	7	1.4	448	2	T12006	NADH2 dehydrogenas	222	7	1.4	1943	2	B64596	toxin-like outer m
150	7	1.4	452	2	T35729	hypothetical prote	223	7	1.4	1969	2	T08875	histidine kinase h
151	7	1.4	454	2	G70797	hypothetical prote	224	7	1.4	2145	2	S61041	glutamate synthase
152	7	1.4	454	2	T26654	hypothetical prote	225	7	1.4	3194	2	D71917	toxin-like outer m
153	7	1.4	455	2	A82957	probable glutamine	226	7	1.4	15281	2	S41309	cyclosporin synthe
154	7	1.4	460	2	AH0968	sodium,galactoside	227	6	1.2	13	2	D61491	seed protein ws-4
155	7	1.4	461	2	D84971	phosphotransferase	228	6	1.2	17	2	A41053	glutamate receptor
156	7	1.4	462	2	B32840	anthranilate synth	229	6	1.2	18	2	A41877	Lcrkc - Yersinia p
157	7	1.4	463	2	B64476	NADP-dependent gly	230	6	1.2	19	2	S69166	ferredoxin b - Jap
158	7	1.4	465	2	I39473	Na+-dependent phos	231	6	1.2	32	2	C61491	seed protein ws-3
159	7	1.4	467	2	A48916	sodium phosphate t	232	6	1.2	35	2	I54460	Oa2 protein - mous
160	7	1.4	472	2	S28286	hypothetical prote	233	6	1.2	39	2	I64027	hypothetical prote
161	7	1.4	473	2	S20612	triacylglycerol li	234	6	1.2	44	2	C35156	dihydrolipoamide d
162	7	1.4	473	2	A38627	gamma-aminobutyric	235	6	1.2	45	2	B35156	dihydrolipoamide d
163	7	1.4	477	2	H83588	probable MFS trans	236	6	1.2	50	2	AE2781	conserved hypotnet
164	7	1.4	479	2	E91195	probable permealase	237	6	1.2	50	2	C97409	hypothetical prote
165	7	1.4	479	2	F86042	probable permealase	238	6	1.2	52	1	S07502	gene 5.9 protein -
166	7	1.4	479	2	C65167	hypothetical 5l.0	239	6	1.2	52	1	S42315	gene 5.9 protein -
167	7	1.4	487	2	F72200	glutamate-tRNA lig	240	6	1.2	55	2	A83500	periplasmic nitrat
168	7	1.4	493	2	G88553	protein C38C10.2 [241	6	1.2	56	2	G72355	hypothetical prote
169	7	1.4	508	2	T05156	probable glucose t	242	6	1.2	57	2	S19088	dihydrolipoamide d
170	7	1.4	515	2	E71447	probable PSII Di p	243	6	1.2	59	2	C81995	hypothetical prote
171	7	1.4	521	2	A53153	glucose transport	244	6	1.2	59	2	F84111	hypothetical prote
172	7	1.4	528	2	AG1937	glucose-6-phosphat	245	6	1.2	60	2	B84811	hypothetical prote
173	7	1.4	531	2	T12406	cytochrome-c oxida	246	6	1.2	61	2	E86084	hypothetical prote
174	7	1.4	539	2	F72288	methyl-accepting c	247	6	1.2	61	2	F96021	hypothetical prote
175	7	1.4	542	1	A70220	phosphotransferase	248	6	1.2	62	2	E89853	hypothetical prote

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:55:49 ; Search time 106.586 Seconds
(without alignments)
1794.174 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 495

Sequence: 1 MKSPVSLAPSGEGSDRT.....LFAKGEVQNAISDHQGRN 495

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1726216 seqs, 386330316 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10A_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	495	100.0	495	9	US-09-776-865-4
2	495	100.0	495	17	US-10-823-506-4
3	61	12.3	194	15	US-10-296-115-1200
4	61	12.3	272	15	US-10-264-237-2351
5	61	12.3	284	15	US-10-264-049-2477
6	61	12.3	495	9	US-09-915-181A-7
7	61	12.3	495	16	US-10-755-889-588
8	61	12.3	495	17	US-10-823-506-2
9	61	12.3	536	9	US-09-776-865-2
10	61	12.3	536	17	US-10-823-506-8
11	14	2.8	495	17	US-10-823-506-10
					Sequence 4, Appli
					Sequence 1200, Ap
					Sequence 2351, Ap
					Sequence 2477, Ap
					Sequence 7, Appli
					Sequence 588, Appl
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 8, Appli
					Sequence 10, Appli

12	14	2.8	495	17	US-10-823-506-12	Sequence 12, Appli
13	12	2.4	78	14	US-10-029-386-27708	Sequence 27708, A
14	11	2.2	578	9	US-09-740-041-4	Sequence 4, Appli
15	11	2.2	578	14	US-10-389-967-4	Sequence 4, Appli
16	11	2.2	582	9	US-09-915-181A-4	Sequence 4, Appli
17	11	2.2	582	15	US-10-205-331-7	Sequence 7, Appli
18	11	2.2	582	16	US-10-734-731-10	Sequence 10, Appli
19	11	2.2	582	16	US-10-734-731-12	Sequence 12, Appli
20	11	2.2	582	16	US-10-734-731-14	Sequence 14, Appli
21	11	2.2	582	16	US-10-807-500-10	Sequence 10, Appli
22	11	2.2	582	16	US-10-807-500-12	Sequence 12, Appli
23	11	2.2	582	16	US-10-807-500-14	Sequence 14, Appli
24	10	2.0	465	9	US-09-915-181A-8	Sequence 8, Appli
25	10	2.0	544	15	US-10-369-493-5873	Sequence 5873, Ap
26	10	2.0	573	15	US-10-369-493-5871	Sequence 5871, Ap
27	10	2.0	573	15	US-10-369-493-5872	Sequence 5872, Ap
28	9	1.8	1803	16	US-10-437-963-188751	Sequence 188751,
29	8	1.6	71	16	US-10-425-115-288336	Sequence 288336,
30	8	1.6	78	16	US-10-437-963-154917	Sequence 154917,
31	8	1.6	93	16	US-10-425-115-366597	Sequence 366597,
32	8	1.6	122	14	US-10-312-187-12	Sequence 12, Appli
33	8	1.6	124	16	US-10-425-115-317259	Sequence 317259,
34	8	1.6	143	16	US-10-425-115-335986	Sequence 335986,
35	8	1.6	232	15	US-10-425-114-43907	Sequence 43907, A
36	8	1.6	235	15	US-10-282-122A-56379	Sequence 56379, A
37	8	1.6	264	15	US-10-108-260A-3281	Sequence 3281, Ap
38	8	1.6	269	16	US-10-437-963-130004	Sequence 130004,
39	8	1.6	276	16	US-10-437-963-138193	Sequence 138193,
40	8	1.6	283	16	US-10-437-963-154806	Sequence 154806,
41	8	1.6	291	16	US-10-437-963-156500	Sequence 156500,
42	8	1.6	337	16	US-10-437-963-195084	Sequence 195084,
43	8	1.6	360	15	US-10-282-122A-45021	Sequence 45021, A
44	8	1.6	365	13	US-10-084-700-32	Sequence 32, Appli
45	8	1.6	365	15	US-10-369-493-1761	Sequence 1761, Ap
46	8	1.6	384	15	US-10-424-599-183068	Sequence 183068,
47	8	1.6	391	16	US-10-437-963-141744	Sequence 141744,
48	8	1.6	448	16	US-10-437-963-125768	Sequence 125768,
49	8	1.6	455	16	US-10-437-963-138196	Sequence 138196,
50	8	1.6	478	16	US-10-437-963-180400	Sequence 180400,
51	8	1.6	479	15	US-10-369-493-23364	Sequence 23364, A
52	8	1.6	484	16	US-10-437-963-138203	Sequence 138203,
53	8	1.6	488	16	US-10-437-963-130220	Sequence 130220,
54	8	1.6	493	16	US-10-437-963-124551	Sequence 124551,
55	8	1.6	495	16	US-10-437-963-141738	Sequence 141738,
56	8	1.6	512	16	US-10-425-115-335979	Sequence 335979,
57	8	1.6	556	16	US-10-437-963-113498	Sequence 113498,
58	8	1.6	560	9	US-09-991-212A-4	Sequence 4, Appli
59	8	1.6	560	10	US-09-915-181A-5	Sequence 5, Appli
60	8	1.6	560	10	US-09-965-522-4	Sequence 4, Appli
61	8	1.6	560	15	US-10-314-790-5	GENERAL INFORMA
62	8	1.6	560	16	US-10-734-731-2	Sequence 2, Appli
63	8	1.6	560	16	US-10-734-731-4	Sequence 4, Appli
64	8	1.6	560	16	US-10-734-731-6	Sequence 6, Appli
65	8	1.6	560	16	US-10-734-731-8	Sequence 8, Appli
66	8	1.6	560	16	US-10-807-500-2	Sequence 2, Appli
67	8	1.6	560	16	US-10-807-500-4	Sequence 4, Appli
68	8	1.6	560	16	US-10-807-500-6	Sequence 6, Appli
69	8	1.6	560	16	US-10-807-500-8	Sequence 8, Appli
70	8	1.6	560	17	US-10-877-818-4	Sequence 188694,
71	8	1.6	561	16	US-10-437-963-188694	Sequence 188694,
72	8	1.6	566	15	US-10-287-226-300	Sequence 300, App
73	8	1.6	580	16	US-10-437-963-146162	Sequence 146162,
74	8	1.6	589	9	US-09-740-041-2	Sequence 2, Appli
75	8	1.6	589	14	US-10-389-967-2	Sequence 2, Appli
76	8	1.6	589	18	US-10-499-731-28	Sequence 28, Appli
77	8	1.6	594	16	US-10-437-963-138200	Sequence 138200,
78	8	1.6	601	18	US-10-499-731-46	Sequence 46, Appli
79	8	1.6	638	16	US-10-437-963-195088	Sequence 195088,
80	8	1.6	642	16	US-10-437-963-164012	Sequence 164012,
81	8	1.6	643	16	US-10-437-963-124148	Sequence 124148,
82	8	1.6	652	16	US-10-437-963-124155	Sequence 124155,
83	8	1.6	656	16	US-10-437-963-141761	Sequence 141761,
84	8	1.6	656	16	US-10-437-963-141769	Sequence 141769,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:46:03 ; Search time 28.807 Seconds
(without alignments)
1282.720 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 495

Sequence: 1 MKSPVSLAPSGEGSDRT.....LFAKGEVQNWAIHQHGRN 495

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pap:*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pap:*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pap:*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pap:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pap:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	495	100.0	495	4	US-09-359-167-4
2	61	12.3	495	4	US-09-359-167-2
3	61	12.3	495	4	US-09-915-181A-7
4	61	12.3	495	4	US-09-359-167-8
5	61	12.3	585	4	US-09-949-016-7705
6	14	2.8	495	4	US-09-359-167-10
7	14	2.8	495	4	US-09-359-167-12
8	12	2.4	436	4	US-09-949-016-11448
9	12	2.4	470	2	US-08-724-394A-10
10	11	2.2	578	4	US-09-740-041-4
11	11	2.2	582	4	US-09-915-181A-4
12	10	2.0	465	4	US-09-915-181A-8
13	8	1.6	140	4	US-09-270-767-33069
14	8	1.6	140	4	US-09-270-767-48286
15	8	1.6	171	4	US-09-902-540-15679
16	8	1.6	234	4	US-09-252-991A-28420
17	8	1.6	367	4	US-09-328-352-6126
18	8	1.6	380	4	US-09-949-016-7053
19	8	1.6	401	4	US-09-328-352-6962
20	8	1.6	497	4	US-09-949-016-6616
21	8	1.6	560	1	US-08-647-484-2
22	8	1.6	560	1	US-08-647-481-2
23	8	1.6	560	1	US-08-430-033A-2
24	8	1.6	560	1	US-08-805-118A-2
25	8	1.6	560	3	US-09-391-958A-4
26	8	1.6	560	4	US-09-915-181A-5
27	8	1.6	560	5	PCT-US96-05792-2

28	8	1.6	567	4	US-09-949-016-11354	Sequence 11354, A
29	8	1.6	589	4	US-09-740-041-2	Sequence 2, Appli
30	8	1.6	850	4	US-09-915-181A-3	Sequence 3, Appli
31	7	1.4	48	4	US-09-270-767-60724	Sequence 60724, A
32	7	1.4	91	4	US-09-543-681A-8118	Sequence 8118, Ap
33	7	1.4	93	4	US-09-248-796A-22842	Sequence 22842, A
34	7	1.4	96	4	US-09-252-991A-16613	Sequence 16613, A
35	7	1.4	102	4	US-09-710-279-2592	Sequence 2592, Ap
36	7	1.4	110	4	US-09-248-796A-14320	Sequence 14320, A
37	7	1.4	122	3	US-09-134-001C-4987	Sequence 4987, Ap
38	7	1.4	135	4	US-09-252-991A-25112	Sequence 25112, A
39	7	1.4	140	1	US-08-133-347-5	Sequence 5, Appli
40	7	1.4	142	4	US-09-949-016-7180	Sequence 7180, Ap
41	7	1.4	157	4	US-09-673-395A-267	Sequence 267, App
42	7	1.4	157	4	US-09-270-767-44201	Sequence 44201, A
43	7	1.4	163	4	US-09-270-767-32926	Sequence 32926, A
44	7	1.4	163	4	US-09-270-767-48143	Sequence 48143, A
45	7	1.4	177	4	US-09-489-039A-11522	Sequence 11522, A
46	7	1.4	187	3	US-09-194-185-1	Sequence 1, Appli
47	7	1.4	195	4	US-09-949-016-10851	Sequence 10851, A
48	7	1.4	204	4	US-09-549-831-6	Sequence 6, Appli
49	7	1.4	212	4	US-09-934-901-10	Sequence 10, Appli
50	7	1.4	212	4	US-09-934-868-20	Sequence 20, Appli
51	7	1.4	212	4	US-09-248-796A-20354	Sequence 20354, A
52	7	1.4	212	4	US-10-321-210-10	Sequence 10, Appli
53	7	1.4	212	4	US-10-320-874-10	Sequence 10, Appli
54	7	1.4	225	4	US-09-252-991A-29348	Sequence 29348, A
55	7	1.4	225	4	US-09-270-767-45231	Sequence 45231, A
56	7	1.4	228	4	US-09-540-236-2798	Sequence 2798, Ap
57	7	1.4	279	3	US-08-549-515-7	Sequence 7, Appli
58	7	1.4	282	4	US-09-583-110-3800	Sequence 3800, Ap
59	7	1.4	287	4	US-09-107-433-4120	Sequence 4120, Ap
60	7	1.4	291	4	US-09-252-991A-28572	Sequence 28572, A
61	7	1.4	297	4	US-09-252-991A-24412	Sequence 24412, A
62	7	1.4	301	4	US-09-902-540-15018	Sequence 15018, A
63	7	1.4	306	4	US-09-134-000C-4155	Sequence 4155, Ap
64	7	1.4	307	4	US-09-522-714-16	Sequence 16, Appli
65	7	1.4	307	4	US-09-489-039A-10116	Sequence 10116, A
66	7	1.4	308	4	US-09-583-110-2950	Sequence 2950, Ap
67	7	1.4	310	4	US-09-543-681A-6311	Sequence 6311, Ap
68	7	1.4	310	4	US-09-248-796A-17322	Sequence 17322, A
69	7	1.4	311	4	US-09-489-039A-9786	Sequence 9786, Ap
70	7	1.4	313	4	US-09-107-433-3357	Sequence 3357, Ap
71	7	1.4	326	4	US-09-949-016-6679	Sequence 6679, Ap
72	7	1.4	347	4	US-09-949-016-11456	Sequence 11456, A
73	7	1.4	348	4	US-09-252-991A-26643	Sequence 26643, A
74	7	1.4	351	2	US-08-843-530B-31	Sequence 31, Appli
75	7	1.4	351	4	US-09-636-728-28	Sequence 28, Appli
76	7	1.4	362	4	US-09-248-796A-26034	Sequence 26034, A
77	7	1.4	381	4	US-09-721-870-28	Sequence 28, Appli
78	7	1.4	431	1	US-08-311-023-2	Sequence 2, Appli
79	7	1.4	441	4	US-09-134-000C-6501	Sequence 6501, Ap
80	7	1.4	451	4	US-09-578-063-43	Sequence 43, Appli
81	7	1.4	460	4	US-09-252-991A-28387	Sequence 28387, A
82	7	1.4	465	4	US-09-252-991A-18003	Sequence 18003, A
83	7	1.4	467	2	US-08-805-118-3	Sequence 3, Appli
84	7	1.4	467	3	US-09-391-958-3	Sequence 3, Appli
85	7	1.4	469	4	US-09-489-039A-10553	Sequence 10553, A
86	7	1.4	469	4	US-09-248-796A-14616	Sequence 14616, A
87	7	1.4	470	4	US-09-578-063-69	Sequence 69, Appli
88	7	1.4	473	4	US-09-411-132A-8	Sequence 8, Appli
89	7	1.4	480	2	US-08-724-394A-9	Sequence 9, Appli
90	7	1.4	488	3	US-09-099-053-2	Sequence 2, Appli
91	7	1.4	491	4	US-09-252-991A-19014	Sequence 19014, A
92	7	1.4	503	4	US-09-489-039A-10722	Sequence 10722, A
93	7	1.4	512	4	US-09-902-540-15789	Sequence 15789, A
94	7	1.4	540	4	US-09-252-991A-22555	Sequence 22555, A
95	7	1.4	563	4	US-09-915-181A-6	Sequence 6, Appli
96	7	1.4	574	4	US-09-248-796A-16849	Sequence 16849, A
97	7	1.4	576	3	US-08-864-785-1	Sequence 1, Appli
98	7	1.4	579	4	US-09-252-991A-29372	Sequence 29372, A
99	7	1.4	599	4	US-09-949-016-7256	Sequence 7256, Ap
100	7	1.4	651	4	US-09-489-039A-7892	Sequence 7892, Ap

101	7	1.4	663	1	US-08-441-139-7	Sequence 7, Appli	174	6	1.2	65	4	US-09-248-796A-22061	Sequence 22061, A
102	7	1.4	678	1	US-08-282-141-2	Sequence 2, Appli	175	6	1.2	66	4	US-09-543-681A-6679	Sequence 6679, Ap
103	7	1.4	678	1	US-08-435-434-2	Sequence 2, Appli	176	6	1.2	66	4	US-09-673-395A-165	Sequence 165, App
104	7	1.4	678	1	US-08-435-436-2	Sequence 2, Appli	177	6	1.2	68	4	US-09-513-999C-4902	Sequence 4902, Ap
105	7	1.4	678	2	US-08-438-863-2	Sequence 2, Appli	178	6	1.2	69	4	US-09-248-796A-14723	Sequence 14723, A
106	7	1.4	678	2	US-08-438-864-2	Sequence 2, Appli	179	6	1.2	70	4	US-09-489-039A-10796	Sequence 10796, A
107	7	1.4	678	2	US-08-438-862-2	Sequence 2, Appli	180	6	1.2	70	4	US-09-621-976-6964	Sequence 6964, Ap
108	7	1.4	678	3	US-08-628-747-2	Sequence 2, Appli	181	6	1.2	72	4	US-09-248-796A-21901	Sequence 21901, A
109	7	1.4	678	3	US-08-402-253-2	Sequence 2, Appli	182	6	1.2	72	4	US-09-248-796A-27413	Sequence 27413, A
110	7	1.4	678	3	US-08-443-866B-2	Sequence 2, Appli	183	6	1.2	73	4	US-09-107-532A-4574	Sequence 4574, Ap
111	7	1.4	722	4	US-09-252-991A-17407	Sequence 17407, A	184	6	1.2	74	4	US-09-621-976-7165	Sequence 7165, Ap
112	7	1.4	755	4	US-09-642-034-5	Sequence 5, Appli	185	6	1.2	74	4	US-09-270-767-40827	Sequence 40827, A
113	7	1.4	832	4	US-09-252-991A-17267	Sequence 17267, A	186	6	1.2	74	4	US-09-270-767-56043	Sequence 56043, A
114	7	1.4	834	4	US-09-404-967C-5	Sequence 5, Appli	187	6	1.2	75	3	US-08-817-811-97	Sequence 97, Appl
115	7	1.4	850	4	US-09-328-352-7660	Sequence 7660, Ap	188	6	1.2	75	4	US-09-788-006-169	Sequence 169, App
116	7	1.4	859	5	PCT-US95-02792-2	Sequence 2, Appli	189	6	1.2	75	4	US-09-902-540-12774	Sequence 12774, A
117	7	1.4	859	5	PCT-US95-02792-2	Sequence 2, Appli	190	6	1.2	75	4	US-09-902-540-13864	Sequence 13864, A
118	7	1.4	877	1	US-08-430-925A-4	Sequence 4, Appli	191	6	1.2	76	3	US-08-817-811-21	Sequence 21, Appl
119	7	1.4	891	4	US-09-949-016-7798	Sequence 7798, Ap	192	6	1.2	77	4	US-09-489-039A-9235	Sequence 9235, Ap
120	7	1.4	937	4	US-09-949-016-8286	Sequence 8286, Ap	193	6	1.2	77	4	US-09-489-039A-10094	Sequence 10094, A
121	7	1.4	947	4	US-09-252-991A-29359	Sequence 29359, A	194	6	1.2	77	4	US-09-489-039A-11231	Sequence 11231, A
122	7	1.4	947	4	US-09-198-452A-478	Sequence 478, App	195	6	1.2	77	4	US-09-489-039A-11344	Sequence 11344, A
123	7	1.4	967	4	US-09-438-185A-453	Sequence 453, App	196	6	1.2	77	4	US-09-489-039A-11503	Sequence 11503, A
124	7	1.4	1046	4	US-09-252-991A-16902	Sequence 16902, A	197	6	1.2	77	4	US-09-489-039A-11942	Sequence 11942, A
125	7	1.4	1055	4	US-09-404-967C-2	Sequence 2, Appli	198	6	1.2	77	4	US-09-489-039A-12092	Sequence 12092, A
126	7	1.4	1075	4	US-09-404-967C-4	Sequence 4, Appli	199	6	1.2	77	4	US-09-489-039A-12620	Sequence 12620, A
127	7	1.4	1121	4	US-09-949-016-11048	Sequence 11048, A	200	6	1.2	77	4	US-09-621-976-7169	Sequence 7169, Ap
128	7	1.4	1169	1	US-08-315-468-4	Sequence 4, Appli	201	6	1.2	77	4	US-09-513-999C-5363	Sequence 5363, Ap
129	7	1.4	1255	4	US-09-902-540-12628	Sequence 12628, A	202	6	1.2	78	4	US-09-749-637A-399	Sequence 399, App
130	7	1.4	1309	4	US-09-975-413A-10	Sequence 10, Appl	203	6	1.2	80	4	US-09-328-352-6720	Sequence 6720, Ap
131	7	1.4	1318	4	US-10-237-551-197	Sequence 197, App	204	6	1.2	80	4	US-09-248-796A-27193	Sequence 27193, A
132	7	1.4	1341	4	US-09-252-991A-26785	Sequence 26785, A	205	6	1.2	81	4	US-09-489-039A-8805	Sequence 8805, Ap
133	7	1.4	1374	4	US-09-900-425B-2	Sequence 2, Appli	206	6	1.2	81	4	US-09-621-976-4256	Sequence 4256, Ap
134	7	1.4	1374	4	US-08-588-985-2	Sequence 2, Appli	207	6	1.2	81	4	US-09-107-433-2805	Sequence 2805, Ap
135	7	1.4	1865	1	US-08-971-988-2	Sequence 2, Appli	208	6	1.2	82	4	US-09-270-767-56738	Sequence 56738, A
136	7	1.4	1865	4	US-09-949-016-6069	Sequence 6069, Ap	209	6	1.2	83	3	US-09-134-001C-4659	Sequence 4659, Ap
137	7	1.4	3647	4	US-09-949-016-10932	Sequence 10932, A	210	6	1.2	83	4	US-09-489-039A-10406	Sequence 10406, A
138	7	1.4	15281	2	US-08-471-119A-2	Sequence 2, Appli	211	6	1.2	83	4	US-09-248-796A-25200	Sequence 25200, A
139	6	1.2	6	5	PCT-US92-01433A-7	Sequence 7, Appli	212	6	1.2	84	4	US-09-489-039A-9870	Sequence 9870, Ap
140	6	1.2	10	4	US-09-653-465B-11	Sequence 11, Appl	213	6	1.2	85	2	US-08-479-078-12	Sequence 12, Appl
141	6	1.2	10	4	US-09-191-593-41	Sequence 41, Appl	214	6	1.2	85	4	US-09-252-991A-19059	Sequence 19059, A
142	6	1.2	12	4	US-10-118-575A-18	Sequence 18, Appl	215	6	1.2	88	4	US-09-248-796A-15960	Sequence 15960, A
143	6	1.2	13	4	US-09-190-976B-1	Sequence 2, Appli	216	6	1.2	89	4	US-09-107-532A-6179	Sequence 6179, Ap
144	6	1.2	16	2	US-08-642-406A-12	Sequence 12, Appl	217	6	1.2	90	4	US-09-270-767-39428	Sequence 39428, A
145	6	1.2	16	3	US-09-199-534-12	Sequence 12, Appl	218	6	1.2	90	4	US-09-270-767-54645	Sequence 54645, A
146	6	1.2	16	4	US-09-199-534-12	Sequence 12, Appl	219	6	1.2	91	4	US-09-270-767-31960	Sequence 31960, A
147	6	1.2	20	3	US-08-744-419-10	Sequence 10, Appl	220	6	1.2	91	4	US-09-270-767-47177	Sequence 47177, A
148	6	1.2	24	2	US-09-047-026A-16	Sequence 16, Appl	221	6	1.2	92	4	US-09-621-976-7014	Sequence 7014, Ap
149	6	1.2	25	4	US-09-270-767-59886	Sequence 59886, A	222	6	1.2	92	4	US-09-513-999C-4421	Sequence 4421, Ap
150	6	1.2	28	4	US-09-091-097-53	Sequence 53, Appl	223	6	1.2	95	4	US-09-543-681A-8178	Sequence 8178, Ap
151	6	1.2	34	1	US-07-745-206A-5	Sequence 5, Appli	224	6	1.2	95	4	US-09-489-039A-9597	Sequence 9597, Ap
152	6	1.2	34	1	US-08-455-543A-50	Sequence 50, Appl	225	6	1.2	95	4	US-09-270-767-40845	Sequence 40845, A
153	6	1.2	34	2	US-08-223-305C-50	Sequence 50, Appl	226	6	1.2	95	4	US-09-270-767-56061	Sequence 56061, A
154	6	1.2	34	2	US-08-311-363-5	Sequence 5, Appli	227	6	1.2	96	4	US-09-746-801A-41	Sequence 41, Appl
155	6	1.2	36	2	US-08-846-762-59	Sequence 59, Appl	228	6	1.2	96	4	US-09-270-767-61787	Sequence 61787, A
156	6	1.2	40	4	US-09-602-777A-12	Sequence 12, Appl	229	6	1.2	96	4	US-09-513-999C-6655	Sequence 6655, Ap
157	6	1.2	41	4	US-09-270-767-36936	Sequence 36936, A	230	6	1.2	98	4	US-09-513-999C-4280	Sequence 4280, Ap
158	6	1.2	41	4	US-09-270-767-52153	Sequence 52153, A	231	6	1.2	99	4	US-09-270-767-59649	Sequence 59649, A
159	6	1.2	42	4	US-09-270-767-31926	Sequence 31926, A	232	6	1.2	100	4	US-09-252-991A-19250	Sequence 19250, A
160	6	1.2	42	4	US-09-270-767-47143	Sequence 47143, A	233	6	1.2	100	4	US-09-513-999C-5398	Sequence 5398, Ap
161	6	1.2	44	4	US-09-205-258-257	Sequence 257, App	234	6	1.2	101	4	US-09-252-991A-25464	Sequence 25464, A
162	6	1.2	45	3	US-09-018-635-39	Sequence 39, Appl	235	6	1.2	101	4	US-09-513-999C-4281	Sequence 4281, Ap
163	6	1.2	45	4	US-09-912-962-39	Sequence 39, Appl	236	6	1.2	101	4	US-09-513-999C-4992	Sequence 4992, Ap
164	6	1.2	50	3	US-09-262-856A-1	Sequence 1, Appli	237	6	1.2	101	4	US-09-513-999C-4993	Sequence 4993, Ap
165	6	1.2	50	4	US-09-471-276-998	Sequence 998, App	238	6	1.2	102	4	US-09-270-767-61824	Sequence 61824, A
166	6	1.2	51	4	US-09-270-767-76746	Sequence 76746, A	239	6	1.2	103	4	US-09-489-039A-7262	Sequence 7262, Ap
167	6	1.2	56	4	US-09-755-665-53	Sequence 53, Appl	240	6	1.2	103	4	US-09-229-304-9	Sequence 9, Appli
168	6	1.2	59	4	US-09-270-767-41846	Sequence 41846, A	241	6	1.2	104	3	US-08-744-419-2	Sequence 2, Appli
169	6	1.2	59	4	US-09-270-767-57088	Sequence 57088, A	242	6	1.2	104	4	US-09-545-894-4	Sequence 4, Appli
170	6	1.2	61	4	US-09-248-796A-24784	Sequence 24784, A	243	6	1.2	104	4	US-09-270-767-57166	Sequence 57166, A
171	6	1.2	62	4	US-09-248-796A-21439	Sequence 21439, A	244	6	1.2	106	4	US-09-270-767-44049	Sequence 44049, A
172	6	1.2	62	4	US-09-248-796A-25193	Sequence 25193, A	245	6	1.2	107	4	US-09-621-976-5202	Sequence 5202, Ap
173	6	1.2	64	4	US-09-248-796A-26590	Sequence 26590, A	246	6	1.2	107	4	US-09-248-796A-22545	Sequence 22545, A

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:29:52 ; Search time 108.026 Seconds

(without alignments)
1772.222 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 495

Sequence: 1 MKSPVSLAPSGEGSGDRT.....LFAKGEVQNWALSDHQGRN 495

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A Geneseq 16Dec04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	495	100.0	495	3	AAY45088 Sheep GBS
2	495	100.0	495	4	AAB06519 Sheep gro
3	61	12.3	124	4	AG81307 Human APP
4	61	12.3	194	4	AAM25685 Human pro
5	61	12.3	272	5	ABB89975 Human pro
6	61	12.3	284	5	ABP41345 Human ova
7	61	12.3	314	4	AG65238 Human sod
8	61	12.3	495	3	AAY45087 Partial h
9	61	12.3	495	4	AB66967 Human AST
10	61	12.3	495	4	AAM38959 Human pol
11	61	12.3	495	8	ADJ75516 Marker ge
12	61	12.3	495	8	ADR14587 Human NF-
13	61	12.3	495	8	ADP25217 PRO polyp
14	61	12.3	536	3	AAY45089 Human GBS
15	61	12.3	536	4	AAB06518 Human gro
16	32	6.5	277	4	AAM93914 Human pol
17	32	6.5	277	8	ADL32036 Human pro
18	32	6.5	309	4	AAM40745 Human pol
19	29	5.9	256	4	ABG02032 Novel hum
20	29	5.9	343	4	ABG12999 Novel hum
21	29	5.9	343	6	ABO00820 Polypepti
22	15	3.0	15	4	AG65239 Human sod
23	14	2.8	495	3	AAY45090 Human/She
24	12	2.4	78	8	ABO54074 Human gen
25	12	2.4	436	2	AAY78919 Human hae

26	12	2.4	436	7	ADF90914	Human hep
27	12	2.4	559	4	ABB58701	Drosophil
28	11	2.2	578	7	ADG88331	Rat trans
29	11	2.2	582	4	AAM79273	Human pro
30	11	2.2	582	6	ABG74796	Rat DNPI
31	11	2.2	582	6	ABG74795	Human DNP
32	11	2.2	582	6	ABG74797	Murine DN
33	11	2.2	582	6	ABM04787	Rat Na-de
34	11	2.2	582	7	ADC15494	Mouse DNP
35	11	2.2	582	7	ADC15492	Rat DNPI
36	11	2.2	582	7	ADC15490	Human DNP
37	11	2.2	582	7	ADD01475	Human VGL
38	10	2.0	387	8	ADP29764	Human sec
39	10	2.0	497	4	ABB70142	Drosophil
40	10	2.0	502	4	ABB65873	Drosophil
41	10	2.0	502	4	ABB60525	Drosophil
42	10	2.0	544	8	ADN23220	Bacterial
43	10	2.0	573	8	ADN23219	Bacterial
44	10	2.0	573	8	ADN23218	Bacterial
45	9	1.8	363	3	AAO8399	Arabidops
46	9	1.8	384	3	AAO8398	Arabidops
47	9	1.8	462	4	ABB70143	Drosophil
48	8	1.6	79	2	RAY59932	Human myo
49	8	1.6	122	5	AAU81759	Partial p
50	8	1.6	234	7	ABO79674	Pseudomon
51	8	1.6	235	6	ABU28455	Protein e
52	8	1.6	264	7	ADM04596	Human pro
53	8	1.6	273	8	ADN46581	Thermococ
54	8	1.6	360	6	ABU17097	Protein e
55	8	1.6	365	6	ABR53883	Protein s
56	8	1.6	365	7	ADK63542	Disease t
57	8	1.6	365	8	ADN13108	Bacterial
58	8	1.6	367	6	ADA34839	Acinetoba
59	8	1.6	401	6	ADA35675	Acinetoba
60	8	1.6	479	8	ADS44934	Bacterial
61	8	1.6	560	2	AAW05148	Human bra
62	8	1.6	560	2	AAW70500	Human sod
63	8	1.6	560	6	ABG74794	Murine BN
64	8	1.6	560	6	ABG74793	Rat BNPI
65	8	1.6	560	6	ABG74792	Human BNP
66	8	1.6	560	6	ABG74791	Human BNP
67	8	1.6	560	7	ADC15488	Mouse BNP
68	8	1.6	560	7	ADC15484	Human BNP
69	8	1.6	560	7	ADC15482	Human BNP
70	8	1.6	560	7	ADC15486	Rat BNPI
71	8	1.6	560	7	ADD01474	Human VGL
72	8	1.6	566	7	ADJ95072	Novel NOV
73	8	1.6	567	4	AAO13870	Human pol
74	8	1.6	589	5	AAU93329	Human tra
75	8	1.6	589	5	AAO30994	Human tra
76	8	1.6	589	7	ADD01392	Human TCH
77	8	1.6	589	7	ADG88329	Human tra
78	8	1.6	589	8	ADR10021	Human pro
79	8	1.6	601	7	ADD01410	Mouse TCH
80	8	1.6	835	5	AAU81754	Human clk
81	8	1.6	837	8	ABM81694	Tumour-as
82	8	1.6	860	5	ABU07689	Rat gluta
83	8	1.6	874	6	ABU65243	Novel hum
84	8	1.6	2015	4	ABB65890	Drosophil
85	8	1.6	4128	7	ADB85120	Mouse DNA
86	7	1.4	14	3	ADC16844	Human sin
87	7	1.4	19	7	ADC39960	Yellow gr
88	7	1.4	22	8	ADJ79763	Subunit I
89	7	1.4	22	8	ADN12017	Subunit I
90	7	1.4	24	7	ADP76655	Novel hum
91	7	1.4	31	4	ABG12078	Novel hum
92	7	1.4	31	4	ABG13776	Novel hum
93	7	1.4	32	2	AAH86113	S. pneumo
94	7	1.4	34	3	AAH38588	Gene 13 h
95	7	1.4	40	5	AAU89268	Insulin/i
96	7	1.4	40	6	ABJ25329	SAC isola
97	7	1.4	44	3	ABJ25329	Human sec
98	7	1.4	44	5	AAU90770	Insulin/i

99	7	1.4	44	6	ABU01493	Abu01493 S. pneumo	172	7	1.4	207	2	AAR05800	Aar05800 Tilapia (
100	7	1.4	76	4	AAU65452	Aau65452 Propionib	173	7	1.4	207	2	AAR08122	Aar08122 Tilapia g
101	7	1.4	76	4	ABM61971	Abm61971 Propionib	174	7	1.4	207	5	ABB89411	Abb89411 Human pol
102	7	1.4	81	5	ABP08020	Abp08020 Human ORF	175	7	1.4	212	5	ABB09165	Abb09165 Methylo
103	7	1.4	91	7	ADP07833	Adp07833 Bacterial	176	7	1.4	212	5	AAE22300	Aae22300 Methylo
104	7	1.4	96	7	ABO67867	AbO67867 Pseudomon	177	7	1.4	212	5	ABG61560	Abg61560 High grow
105	7	1.4	99	4	AAU30972	Aau30972 Novel hum	178	7	1.4	212	5	ABU05797	Abu05797 M. tuberc
106	7	1.4	102	4	AAQ82749	Aaq82749 S. epider	179	7	1.4	222	6	ABM68944	Abm68944 Phototrab
107	7	1.4	103	6	ADA54545	Ada54545 Human tra	180	7	1.4	222	6	AAO15070	Aao15070 Humicola
108	7	1.4	107	5	ABP35197	Abp35197 Human tra	181	7	1.4	223	5	AAO15070	Aao15070 Humicola
109	7	1.4	110	4	AAQ91463	Aaq91463 Human imm	182	7	1.4	223	5	AAQ80602	Aaq80602 Humicola
110	7	1.4	115	5	ABP31660	Abp31660 Human ORF	183	7	1.4	225	7	ABO80602	AbO80602 Pseudomon
111	7	1.4	115	8	ADR66927	Adr66927 Human pro	184	7	1.4	227	4	ABG16682	Abg16682 Novel hum
112	7	1.4	115	8	ADR66029	Adr66029 Human pro	185	7	1.4	228	5	AAO15062	Aao15062 Endogluca
113	7	1.4	122	5	ABP40142	Abp40142 Staphyloc	186	7	1.4	228	5	ADL05112	Adl05112 M. catarr
114	7	1.4	122	8	ADSO7186	Adso7186 Staphyloc	187	7	1.4	231	4	ABB70615	Abb70615 Drosophil
115	7	1.4	126	6	ABU49557	Abu49557 Protein e	188	7	1.4	231	8	ADN48262	Adn48262 Thermococ
116	7	1.4	132	4	ABM59519	Abm59519 Human sec	189	7	1.4	264	8	ADI19308	Adi19308 Ag2/PRAL-
117	7	1.4	132	7	ADM25718	Adm25718 Hyperther	190	7	1.4	266	6	ABP75751	Abp75751 Human sec
118	7	1.4	133	4	ABE68834	AbE68834 Drosophil	191	7	1.4	270	6	ADQ90464	Adq90464 T tengcon
119	7	1.4	135	7	ABO76366	AbO76366 Pseudomon	192	7	1.4	270	8	ADH45300	Adh45300 Human enz
120	7	1.4	136	7	ADM04743	Adm04743 Human pro	193	7	1.4	279	3	AAV88627	Aav88627 H. influe
121	7	1.4	136	8	ADQ96242	Adq96242 T cell ac	194	7	1.4	279	5	ABU05927	Abu05927 M. tuberc
122	7	1.4	145	2	AAW94926	Aaw94926 Human phe	195	7	1.4	282	6	ABU01530	Abu01530 S. pneumo
123	7	1.4	146	8	ADS17938	AdS17938 Coccidioi	196	7	1.4	282	6	ABP81392	Abp81392 Streptoco
124	7	1.4	147	4	AAW43546	Aaw43546 Human pol	197	7	1.4	282	6	ABU44997	Abu44997 Protein e
125	7	1.4	147	4	AAU19939	Aau19939 Novel hum	198	7	1.4	282	6	ADK47285	Adk47285 Streptoco
126	7	1.4	147	4	AAU87555	Aau87555 Novel cen	199	7	1.4	282	8	ADM92159	Adm92159 S. pneumon
127	7	1.4	147	8	ADI54870	Adi54870 Novel hum	200	7	1.4	284	4	AAU38138	Aau38138 Salmonell
128	7	1.4	147	8	ADM24567	Adm24567 Human PRO	201	7	1.4	284	4	AAU34730	Aau34730 E. coli c
129	7	1.4	155	3	AAE53291	Aae53291 Human col	202	7	1.4	284	6	ABM68248	Abm68248 Phototrab
130	7	1.4	155	3	AAE24501	Aae24501 Human sec	203	7	1.4	284	6	ABU50326	Abu50326 Protein e
131	7	1.4	155	4	AAQ74756	Aaq74756 Human col	204	7	1.4	284	6	ABU28786	Abu28786 Protein e
132	7	1.4	157	2	AAV60067	Aav60067 Human end	205	7	1.4	284	6	ABU40658	Abu40658 Protein e
133	7	1.4	161	6	ABU20231	Abu20231 Protein e	206	7	1.4	284	6	ABU47130	Abu47130 Protein e
134	7	1.4	163	3	AAQ13183	Aaq13183 Arabidops	207	7	1.4	284	6	ABU28308	Abu28308 Protein e
135	7	1.4	166	8	ADS17940	AdS17940 Coccidioi	208	7	1.4	284	6	ABU31668	Abu31668 Protein e
136	7	1.4	171	3	AAE13182	Aae13182 Arabidops	209	7	1.4	284	6	ABU47728	Abu47728 Protein e
137	7	1.4	172	3	AAE56835	Aae56835 Human pro	210	7	1.4	287	8	ADR95485	Adr95485 Novel S.
138	7	1.4	172	6	ABG99987	Abg99987 Human nov	211	7	1.4	289	4	AAU35600	Aau35600 Haemophil
139	7	1.4	173	7	ADM26716	Adm26716 Hyperther	212	7	1.4	289	6	ABU30505	Abu30505 Protein e
140	7	1.4	175	6	ABP78104	Abp78104 N. gonorr	213	7	1.4	291	7	ABO79826	AbO79826 Pseudomon
141	7	1.4	177	7	ABO65005	AbO65005 Klebsiell	214	7	1.4	294	2	AAW70319	Aaw70319 Secreted
142	7	1.4	185	1	AAAP90694	Aap90694 Sequence	215	7	1.4	294	8	ABO84924	AbO84924 Human can
143	7	1.4	185	7	ADC39950	Adc39950 Yellow gr	216	7	1.4	295	2	AAV25758	Aav25758 Human sec
144	7	1.4	186	1	AAQ90616	Aaq90616 Fish grow	217	7	1.4	297	7	ABO75666	AbO75666 Pseudomon
145	7	1.4	187	1	AAE81244	Aae81244 Sequence	218	7	1.4	301	8	ADN47594	Adn47594 Thermococ
146	7	1.4	187	2	AAW71372	Aaw71372 Synthetic	219	7	1.4	302	8	ADL57159	Adl57159 Human NOV
147	7	1.4	187	2	AAW49019	Aaw49019 Korean ro	220	7	1.4	303	6	ABU24928	Abu24928 Protein e
148	7	1.4	188	2	AAE07353	Aae07353 Fish grow	221	7	1.4	304	4	AAE96280	Aae96280 Putative
149	7	1.4	188	2	AAE24137	Aae24137 Fish grow	222	7	1.4	304	5	ABP30465	Abp30465 Streptoco
150	7	1.4	192	4	AAE61301	Aae61301 Human tra	223	7	1.4	304	7	ADM05529	Adm05529 Human pro
151	7	1.4	192	5	ABO09599	AbO09599 Human pal	224	7	1.4	306	5	ABP27232	Abp27232 Streptoco
152	7	1.4	192	8	ADQ96240	Adq96240 T cell ac	225	7	1.4	306	7	ADH86270	Adh86270 Enterococ
153	7	1.4	192	8	ADQ96136	Adq96136 T cell ac	226	7	1.4	307	3	AAE18901	Aae18901 A maize c
154	7	1.4	200	1	AAAP90732	Aap90732 Fish grow	227	7	1.4	307	4	ABG12674	Abg12674 Novel hum
155	7	1.4	200	4	AAE59504	Aae59504 Human sec	228	7	1.4	307	7	ABO63599	AbO63599 Klebsiell
156	7	1.4	200	7	ADC86463	Adc86463 Human GPC	229	7	1.4	308	4	AAO1075	Aao1075 CFE 78 pr
157	7	1.4	201	4	ABG13795	Abg13795 Novel hum	230	7	1.4	308	5	ABP25574	Abp25574 Streptoco
158	7	1.4	202	4	ABE63463	AbE63463 Drosophil	231	7	1.4	308	6	ABU02385	Abu02385 S. pneumo
159	7	1.4	203	1	AAAP90617	Aap90617 Fish grow	232	7	1.4	308	6	ABP81461	Abp81461 Streptoco
160	7	1.4	203	2	AAU07248	Aau07248 Fish grow	233	7	1.4	308	8	ADK46435	Adk46435 Streptoco
161	7	1.4	203	3	AAU77282	Aau77282 Spotted f	234	7	1.4	308	8	ADM92229	Adm92229 S. pneumon
162	7	1.4	204	1	AAE95638	Aae95638 Fish grow	235	7	1.4	308	8	ADR83899	Adr83899 S. pyogen
163	7	1.4	204	2	AAE10912	Aae10912 Fish grow	236	7	1.4	310	7	ADF06026	Adf06026 Bacterial
164	7	1.4	204	2	AAW27340	Aaw27340 Yellow ta	237	7	1.4	311	7	ABO63269	AbO63269 Klebsiell
165	7	1.4	204	2	AAW25120	Aaw25120 Yellowtai	238	7	1.4	313	8	ADR94722	Adr94722 Novel S.
166	7	1.4	204	2	AAW49020	Aaw49020 Korean ro	239	7	1.4	315	8	ADR09349	Adr09349 Human pro
167	7	1.4	204	5	ABG31737	Abg31737 Yellow gr	240	7	1.4	323	8	ADS43275	Ads43275 Bacterial
168	7	1.4	204	5	ABG31734	Abg31734 Yellow gr	241	7	1.4	326	4	AAU34980	Aau34980 Enterococ
169	7	1.4	204	7	ABE64099	AbE64099 Human pro	242	7	1.4	326	6	ABO07155	AbO07155 Novel hum
170	7	1.4	204	7	ADC50041	Adc50041 Orange-sp	243	7	1.4	326	6	ABU29293	Abu29293 Protein e
171	7	1.4	204	7	ADC39941	Adc39941 Yellow gr	244	7	1.4	326	7	ADF90912	Adf90912 Human hep

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:30:47 ; Search time 119.573 Seconds
(without alignments)
2295.451 Million cell updates/sec

Title: US-09-776-865-2

Perfect score: 536

Sequence: 1: MAAGANTPPRPVQPARPGGF.....LFAKGEVQNALNDHGHGRH 536

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : UniProt_03.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	495	92.4	495	1	S175 HUMAN	Q9nra2 homo sapien
2	61	11.4	495	1	S175 SHEEP	Q9mzd1 ovis aries
3	34	6.3	495	1	S175 MOUSE	Q8bn82 mus musculus
4	11	2.1	495	1	S175 MOUSE	Q96kl7 homo sapien
5	11	2.1	465	1	NPT1 HUMAN	Q14916 homo sapien
6	11	2.1	467	1	Q9H531	Q9h531 homo sapien
7	11	2.1	582	2	Q9P2U8	Q9p2u8 homo sapien
8	11	2.1	582	2	Q920B7	Q920b7 mus musculus
9	11	2.1	582	2	Q8BLE7	Q8ble7 mus musculus
10	11	2.1	582	2	Q9J112	Q9j112 rattus norv
11	10	1.9	172	2	Q7PWK3	Q7pwk3 anopheles g
12	10	1.9	363	2	Q8BWS1	Q8bws1 mus musculus
13	10	1.9	436	1	NPT3 HUMAN	Q00624 homo sapien
14	10	1.9	447	2	Q8VCX8	Q8vcx8 mus musculus
15	10	1.9	466	2	Q96LH1	Q96lh1 homo sapien
16	10	1.9	535	2	Q7PWK4	Q7pwk4 anopheles g
17	10	1.9	544	1	YRT3 CAEEL	Q10046 caenorhabdi
18	10	1.9	573	2	Q09932	Q09932 caenorhabdi
19	9	1.7	69	2	Q91Y99	Q91y99 mus musculus
20	9	1.7	127	2	Q71118	Q71118 trichoplusi
21	9	1.7	286	2	Q7PRG4	Q7prg4 anopheles g
22	9	1.7	390	2	Q921B2	Q921b2 mus musculus
23	9	1.7	465	1	NPT1 MOUSE	Q61983 mus musculus
24	9	1.7	465	1	NPT1 RAT	Q62795 rattus norv
25	9	1.7	465	2	Q8K3H3	Q8k3h3 rattus norv
26	9	1.7	465	2	Q91WV5	Q91wv5 mus musculus
27	9	1.7	465	2	Q6AZ46	Q6az46 rattus norv
28	9	1.7	469	2	Q8CJH8	Q8cjh8 rattus norv
29	9	1.7	475	2	Q6AZ69	Q6az69 rattus norv
30	9	1.7	497	2	Q8MRP7	Q8mrp7 drosophila
31	9	1.7	498	2	Q8CJH9	Q8cjh9 rattus norv

32	9	1.7	715	2	O83248	O83248 treponema p
33	8	1.5	10	2	O7S377	O7s377 neurospora
34	8	1.5	126	2	Q9LRG3	Q9lrg3 chlorella v
35	8	1.5	250	1	HXB9_MOUSE	P20615 mus musculus
36	8	1.5	255	2	O99909	O99909 potamoxygo
37	8	1.5	296	2	Q97CJ3	Q97cj3 thermoplas
38	8	1.5	308	1	LIVH_SALTY	P30295 salmonella
39	8	1.5	342	2	Q21438	Q21438 caenorhabdi
40	8	1.5	369	2	Q72KC7	Q72kc7 thermus the
41	8	1.5	373	2	O53550	O53550 mycobacteri
42	8	1.5	373	2	Q7TWC6	Q7twc6 mycobacteri
43	8	1.5	376	2	Q674M9	Q674m9 trialeurode
44	8	1.5	398	2	O9ZDV1	O9zdv1 rickettsia
45	8	1.5	402	2	Q7TQH3	Q7tqh3 mus musculu
46	8	1.5	465	1	NET1_RABIT	Q28722 o renal sod
47	8	1.5	470	2	Q7YTZ7	Q7ytz7 drosophila
48	8	1.5	476	2	Q8ZZR7	Q8zzr7 pyrobaculum
49	8	1.5	476	2	Q7Q579	Q7q579 anopheles g
50	8	1.5	479	1	DBPA_BACSU	P42305 bacillus su
51	8	1.5	479	2	Q23514	Q23514 caenorhabdi
52	8	1.5	497	2	Q9YZC5	Q9yzc5 homo sapien
53	8	1.5	497	2	Q6NR28	Q6nr28 drosophila
54	8	1.5	497	2	Q7KRJ1	Q7krj1 drosophila
55	8	1.5	498	2	Q8VCY5	Q8vcy5 mus musculu
56	8	1.5	502	2	Q9VDM0	Q9vdm0 drosophila
57	8	1.5	515	2	Q8JFT2	Q8jft2 brachydanio
58	8	1.5	533	2	Q7Q1S5	Q7q1s5 anopheles g
59	8	1.5	533	2	Q9SY99	Q9sy99 arabidopsis
60	8	1.5	544	2	Q7Q580	Q7q580 anopheles g
61	8	1.5	550	2	Q8W033	Q8w033 arabidopsis
62	8	1.5	550	2	Q940H4	Q940h4 arabidopsis
63	8	1.5	558	2	Q8S068	Q8s068 oryza sativ
64	8	1.5	559	2	Q9VYG7	Q9vyg7 drosophila
65	8	1.5	560	2	Q6PCD0	Q6pcd0 homo sapien
66	8	1.5	560	2	Q9P2U7	Q9p2u7 homo sapien
67	8	1.5	560	2	Q62634	Q62634 rattus norv
68	8	1.5	576	2	Q6INC8	Q6inc8 xenopus lae
69	8	1.5	584	2	Q8AW47	Q8aw47 brachydanio
70	8	1.5	588	2	Q8K1Q1	Q8k1q1 rattus norv
71	8	1.5	588	2	Q7TSF2	Q7tsf2 rattus norv
72	8	1.5	589	2	Q8NDX2	Q8ndx2 homo sapien
73	8	1.5	592	2	Q7Q3R3	Q7q3r3 anopheles g
74	8	1.5	601	2	Q8BFU8	Q8bfu8 mus musculu
75	8	1.5	695	2	Q76HF4	Q76hf4 seriola qui
76	8	1.5	728	2	Q9U2E0	Q9u2e0 caenorhabdi
77	8	1.5	863	2	Q9ADL8	Q9adl8 polyangium
78	8	1.5	872	2	Q7SX79	Q7sx79 brachydanio
79	8	1.5	887	2	Q6BCK3	Q6bck3 fugu rubrip
80	8	1.5	888	2	Q6DEI2	Q6dei2 brachydanio
81	8	1.5	888	2	Q7ZZS7	Q7zzs7 brachydanio
82	8	1.5	897	2	Q7ZT42	Q7zt42 brachydanio
83	8	1.5	911	2	Q6BCK4	Q6bck4 fugu rubrip
84	8	1.5	912	2	Q7ZT45	Q7zt45 seriola qui
85	8	1.5	935	2	Q7Q6T9	Q7q6t9 anopheles g
86	8	1.5	940	2	Q7XTN8	Q7xtn8 oryza sativ
87	8	1.5	949	2	Q9SMA4	Q9sma4 oryza sativ
88	8	1.5	953	2	Q9V701	Q9v701 drosophila
89	8	1.5	1349	2	Q9L096	Q9l096 streptomyce
90	8	1.5	2180	2	Q9VAS8	Q9vas8 drosophila
91	8	1.5	4128	1	PRKD_MOUSE	P97313 mus musculu
92	7	1.3	39	2	Q33434	Q33434 equus burch
93	7	1.3	39	2	Q33435	Q33435 equus cabal
94	7	1.3	39	2	Q34433	Q34433 equus quagg
95	7	1.3	50	2	Q6EAN8	Q6ean8 sus scrofa
96	7	1.3	54	2	Q88MK7	Q88mk7 pseudomonas
97	7	1.3	68	2	Q6IG36	Q6ig36 drosophila
98	7	1.3	71	2	Q8H710	Q8h710 oryza sativ
99	7	1.3	73	1	CATZ_BOVIN	P05689 bos taurus
100	7	1.3	73	2	Q96PF5	Q96pf5 homo sapien
101	7	1.3	81	2	Q71I23	Q71i23 lactobacill
102	7	1.3	84	2	Q7U5U8	Q7u5u8 synechococc
103	7	1.3	95	1	KRPT_ANAPL	P08335 anas platyr
104	7	1.3	95	2	Q6UXF0	Q6uxp0 homo sapien

105	7	1.3	99	2	Q924W3	Q924W3 rattus norv	178	7	1.3	187	2	Q7ZV4	Q7zyv4 odontesthes
106	7	1.3	102	2	Q9B219	Q9B219 branchiosteo	179	7	1.3	188	2	Q7PWJ4	Q7pwj4 anopheles g
107	7	1.3	103	2	Q64QW1	Q64qw1 bacterioides	180	7	1.3	188	2	Q6ESW6	Q6esw6 myotis myst
108	7	1.3	104	1	VATF THEVO	Q97cq1 thermoplasm	181	7	1.3	191	1	PMVK HUMAN	Q91126 homo sapien
109	7	1.3	106	2	Q725U9	Q725u9 desulfovibr	182	7	1.3	191	1	PMVK_MOUSE	Q9d1g2 mus musculu
110	7	1.3	107	2	Q13532	Q13532 saccharomyc	183	7	1.3	192	2	Q8YFN3	Q8yfn3 brucella me
111	7	1.3	109	2	Q87TK8	Q87tk8 vibrio para	184	7	1.3	192	2	Q8G280	Q8g280 brucella su
112	7	1.3	114	2	Q32461	Q32461 actinomadr	185	7	1.3	193	2	Q8LW42	Q8lww42 coregonus l
113	7	1.3	115	2	Q37739	Q37739 felis silve	186	7	1.3	193	2	Q8MOM2	Q8m0m2 coregonus l
114	7	1.3	116	2	Q8HCW7	Q8hcv7 plecotus ma	187	7	1.3	193	2	Q8MOM3	Q8m0m3 coregonus l
115	7	1.3	116	2	Q94JE6	Q94je6 oryza sativ	188	7	1.3	193	2	Q8MOM4	Q8m0m4 coregonus l
116	7	1.3	118	2	Q64QD9	Q64qd9 bacterioides	189	7	1.3	193	2	Q8MOM5	Q8m0m5 coregonus l
117	7	1.3	118	2	Q8CRD8	Q8crd8 staphylococ	190	7	1.3	193	2	Q8MOM6	Q8m0m6 coregonus l
118	7	1.3	124	2	Q9T4W9	Q9t4w9 arabacia lix	191	7	1.3	193	2	Q8C4W2	Q8c4w2 mus musculu
119	7	1.3	125	1	NULM ARBLI	Q33756 arabacia lix	192	7	1.3	195	2	Q9Z8B8	Q9z8b8 chlamydia p
120	7	1.3	125	2	Q6M0T0	Q6m0t0 methanococc	193	7	1.3	195	2	Q8JM30	Q8jm30 mamestra co
121	7	1.3	126	1	ACPS VIBCH	Q9kpb6 vibrio chol	194	7	1.3	195	2	Q8QLA9	Q8qla9 mamestra co
122	7	1.3	126	2	QYTH7	Q9yth7 caenorhabdi	195	7	1.3	195	2	Q7IA87	Q7ia87 mamestra co
123	7	1.3	127	2	Q3ER8	Q3er8 bacillus ce	196	7	1.3	195	2	Q8JIN3	Q8jin3 hippoglossu
124	7	1.3	129	1	PMVK PIG	Q29081 sus scrofa	197	7	1.3	196	1	SOMA_FUGRU	Q12980 fugu rubrip
125	7	1.3	129	2	Q8QNH2	Q8qnh2 ectocarpus	198	7	1.3	196	1	SOMA_SIGGU	Q9ibes5 siganus gut
126	7	1.3	131	2	Q7QX3	Q7qxe3 giardia lam	199	7	1.3	196	2	Q823F1	Q823p1 chlamydomph
127	7	1.3	132	2	Q8TVH5	Q8tyh5 methanopyru	200	7	1.3	197	1	Y089 MYCTU	P65346 mycobacteri
128	7	1.3	133	2	Q9Y928	Q9y928 acroporum p	201	7	1.3	197	1	Y092 MYCBO	P65347 mycobacteri
129	7	1.3	135	2	Q8W128	Q8w128 capsicum ex	202	7	1.3	199	2	Q6MKU1	Q6mkul bdellovibri
130	7	1.3	136	2	Q7UX50	Q7ux50 rhodopirell	203	7	1.3	200	2	Q8NGA3	Q8nga3 homo sapien
131	7	1.3	138	1	YCM2 SCHPO	Q94249 schizosacch	204	7	1.3	200	2	Q79237	Q79237 thymallus a
132	7	1.3	142	2	Q7X8R7	Q7x8r7 oryza sativ	205	7	1.3	200	2	Q9T4G7	Q9t4g7 thymallus a
133	7	1.3	143	2	Q91I70	Q9i170 helicoverpa	206	7	1.3	200	2	Q892V5	Q892v5 clostriedium
134	7	1.3	146	2	Q7NGX7	Q7ngx7 gloebacter	207	7	1.3	201	2	Q8JZR7	Q8jzr7 mus musculu
135	7	1.3	148	2	Q9ITJ4	Q9itj4 tupaiid her	208	7	1.3	202	2	Q9NTR6	Q9ntr6 homo sapien
136	7	1.3	150	2	Q94WB2	Q94wb2 ptereleotri	209	7	1.3	202	2	Q9V8V0	Q9v8v0 drosophila
137	7	1.3	150	2	Q94WD3	Q94wd3 ptereleotri	210	7	1.3	202	2	Q9FHR1	Q9fhr1 arabidopsis
138	7	1.3	151	2	Q7PI79	Q7pi79 anopheles g	211	7	1.3	203	1	SOMA_PAGMA	P08591 pagrus majo
139	7	1.3	151	2	Q48800	Q48800 legionella	212	7	1.3	203	1	SOMA_VERVA	Q93566 verasper va
140	7	1.3	151	2	Q6QF8	Q6qf8 legionella	213	7	1.3	203	2	Q7MQW7	Q7mqw7 wolinnella s
141	7	1.3	151	2	Q6QF1	Q6qf1 legionella	214	7	1.3	203	2	Q8JIN5	Q8jin5 cottus kazi
142	7	1.3	151	2	Q6QF3	Q6qf3 legionella	215	7	1.3	204	1	SOMA_ACABU	Q01282 acanthopagr
143	7	1.3	151	2	Q6QFH0	Q6qfh0 legionella	216	7	1.3	204	1	SOMA_ACALA	P45654 acanthopagr
144	7	1.3	152	2	Q02116	Q02116 rhizopus st	217	7	1.3	204	1	SOMA_DICLA	Q05163 dicentrarch
145	7	1.3	153	2	Q9KUY1	Q9kuy1 vibrio chol	218	7	1.3	204	1	SOMA_LANCA	Q01283 lates calca
146	7	1.3	154	2	Q7S1E3	Q7s1e3 neurospora	219	7	1.3	204	1	SOMA_MORSA	P48248 morone saxa
147	7	1.3	155	2	Q8W129	Q8w129 capsicum ca	220	7	1.3	204	1	SOMA_ODOAR	Q91915 odontesthes
148	7	1.3	155	2	Q91459	Q91459 seriola qui	221	7	1.3	204	1	SOMA_OREMO	P34746 oreochromis
149	7	1.3	158	2	Q9CCU7	Q9ccu7 mycobacteri	222	7	1.3	204	1	SOMA_ORENI	P13391 oreochromis
150	7	1.3	160	2	Q9CSW6	Q9csw6 mus musculu	223	7	1.3	204	1	SOMA_PBRFV	Q9d5v3 perca flave
151	7	1.3	161	2	Q6GKY2	Q6gky2 arabidopsis	224	7	1.3	204	1	SOMA_PSECR	Q9i9n4 pseudosciae
152	7	1.3	165	2	Q8UDC2	Q8udc2 agrobacteri	225	7	1.3	204	1	SOMA_SCIOC	Q9ib11 sciaenops o
153	7	1.3	166	2	Q8LXA6	Q8lxa6 debaryomyce	226	7	1.3	204	1	SOMA_SEBSC	P87391 sebastes sc
154	7	1.3	168	2	Q9UTU7	Q9utu7 schizosacch	227	7	1.3	204	1	SOMA_SERQU	P09539 seriola qui
155	7	1.3	168	2	Q7U595	Q7u595 synechococc	228	7	1.3	204	1	SOMA_SPAAU	P29971 sparus aura
156	7	1.3	169	2	Q07698	Q07698 mycobacteri	229	7	1.3	204	1	SOMA_THUTH	P09113 thunnus thy
157	7	1.3	171	2	Q937C3	Q937c3 enterococcu	230	7	1.3	204	1	SOMA_TRITC	Q98uf6 trichogate
158	7	1.3	171	2	Q9KYH0	Q9kyh0 streptomyce	231	7	1.3	204	2	Q8UW02	Q8uw02 epinephelus
159	7	1.3	173	2	Q82HK8	Q82hk8 streptomyce	232	7	1.3	204	2	Q90VV6	Q90vv6 epinephelus
160	7	1.3	176	2	Q9RPS9	Q9rps9 enterococcu	233	7	1.3	204	2	Q90YK4	Q90yk4 sparus aura
161	7	1.3	176	2	Q6N353	Q6n353 rhodopsendo	234	7	1.3	204	2	Q6LAL0	Q6lalo oreochromis
162	7	1.3	179	2	Q7NH52	Q7nh52 gloebacter	235	7	1.3	204	2	Q6QRJ2	Q6qrp2 leptomis cya
163	7	1.3	181	2	Q88ZC1	Q88zc1 lactobacill	236	7	1.3	204	2	Q7SYG9	Q7syg9 siniperca k
164	7	1.3	182	2	Q6EGY7	Q6egy7 plecotus au	237	7	1.3	204	2	Q7TIG3	Q7tig3 epinephelus
165	7	1.3	182	2	Q6EGY8	Q6egy8 plecotus au	238	7	1.3	204	2	Q7T231	Q7t231 monopteris
166	7	1.3	182	2	Q6EGY9	Q6egy9 plecotus ma	239	7	1.3	208	2	Q8STK6	Q8stk6 encephalito
167	7	1.3	182	2	Q6EGZ0	Q6egz0 plecotus ma	240	7	1.3	210	2	Q9RCV9	Q9rcv9 streptomyce
168	7	1.3	182	2	Q6EGZ1	Q6egz1 plecotus au	241	7	1.3	210	2	Q9CXB8	Q9cxb8 mus musculu
169	7	1.3	182	2	Q84455	Q84455 paramacium	242	7	1.3	211	2	Q943X3	Q943x3 oryza sativ
170	7	1.3	185	1	SOMA_KATPE	P20391 katsuwonus	243	7	1.3	211	2	Q7N6C9	Q7n6c9 photorhabdu
171	7	1.3	185	2	Q89NE1	Q89ne1 bradyrhizob	244	7	1.3	213	2	Q8WB47	Q8wb47 bufo macula
172	7	1.3	185	2	Q6UAM5	Q6uam5 tetraodon n	245	7	1.3	213	2	Q6VZK8	Q6vzk8 canaryopus v
173	7	1.3	186	2	Q85609	Q85609 escherichia	246	7	1.3	215	2	Q8RTT8	Q8rtt8 uncultured
174	7	1.3	187	1	SOMA_THUAL	P34747 thunnus alb	247	7	1.3	215	2	Q6SPD5	Q6spd5 uncultured
175	7	1.3	187	2	Q91077	Q91077 lateolabrax	248	7	1.3	217	1	ATOD_HAEIN	P44875 haemophilus
176	7	1.3	187	2	Q91449	Q91449 seriola dum	249	7	1.3	217	1	OPBB_BACSU	Q34561 bacillus su
177	7	1.3	187	2	Q91454	Q91454 sebaetiscus	250	7	1.3	217	1	OPCB_BACSU	Q34878 bacillus su

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:34:43 ; Search time 27.0339 Seconds
(without alignments)
1907.682 Million cell updates/sec

Title: US-09-776-865-2
Perfect score: 536
Sequence: 1 MAGAMTPRPVQPARPGF.....LFAKGEVQNWALNDHGHHRH 536

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR 79:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	2.1	465	2 139473	Na+-dependent phos
2	11	2.1	467	2 A48916	sodium phosphate t
3	10	1.9	544	2 T24633	hypothetical prote
4	10	1.9	573	2 T23589	hypothetical prote
5	9	1.7	465	2 S89915	sodium-phosphate t
6	9	1.7	715	2 H71352	probable sigma fac
7	8	1.5	250	2 A56572	homeotic protein H
8	8	1.5	308	2 AG0992	high-affinity bran
9	8	1.5	308	2 JH0668	high-affinity bran
10	8	1.5	373	2 C70806	probable fadE27 pr
11	8	1.5	396	2 T23619	hypothetical prote
12	8	1.5	398	2 G71733	hypothetical prote
13	8	1.5	465	2 A56410	sodium/phosphate t
14	8	1.5	479	2 E69613	ATP-dependent RNA
15	8	1.5	533	2 T04770	aldehyde dehydroge
16	8	1.5	560	2 I59302	brain specific Na+
17	8	1.5	4128	2 JC6306	protein kinase (EC
18	7	1.3	33	2 S13863	dihydrolipoamide d
19	7	1.3	73	2 A29172	cathepsin B-like c
20	7	1.3	95	1 KRDKF4	keratin B-4, feath
21	7	1.3	107	2 S69295	probable membrane
22	7	1.3	114	2 JC5860	polyketide synthas
23	7	1.3	126	2 F82072	holo-(acyl-carrier
24	7	1.3	133	2 H72476	hypothetical prote
25	7	1.3	138	2 T40963	very hypothetical
26	7	1.3	151	2 S61384	icmW protein - leg
27	7	1.3	152	2 PT0433	progesterone 11alp
28	7	1.3	153	2 E82330	conserved hypothet
29	7	1.3	158	2 C86961	probable tRNA/rRNA

30	1.3	165	2	C97624	petP protein (impo
31	1.3	165	2	AD2847	transcription regu
32	1.3	182	2	T17625	hypothetical prote
33	1.3	185	2	JK0021	somatotropin - ski
34	1.3	187	2	A61123	somatotropin - Moz
35	1.3	187	2	PC1082	somatotropin - Jap
36	1.3	187	2	JU0030	somatotropin - yel
37	1.3	187	2	PC1087	somatotropin - dus
38	1.3	187	2	PC1088	somatotropin - gre
39	1.3	192	2	AI3437	colicin v producti
40	1.3	195	2	G86543	CR276 hypothetical
41	1.3	195	2	C72081	conserved hypothet
42	1.3	197	2	A70750	hypothetical prote
43	1.3	203	2	S00747	somatotropin precu
44	1.3	204	1	STFI	somatotropin precu
45	1.3	204	2	A56642	somatotropin precu
46	1.3	204	2	I51289	somatotropin - Mor
47	1.3	204	2	JN0484	somatotropin precu
48	1.3	204	2	JH0577	somatotropin precu
49	1.3	204	2	A56904	somatotropin precu
50	1.3	204	2	S30451	somatotropin - Aca
51	1.3	204	2	JC4261	somatotropin precu
52	1.3	204	2	S01746	somatotropin precu
53	1.3	217	2	B64092	butyrate-acetoacet
54	1.3	217	2	D69670	glycine betaine/ca
55	1.3	217	2	H69669	choline ABC transp
56	1.3	218	2	A95295	probable ABC trans
57	1.3	225	2	B83083	conserved hypothet
58	1.3	232	2	AF3406	hypothetical expor
59	1.3	233	2	A69004	CDPdiacylglycerol-
60	1.3	236	2	S42069	TEGT protein - rat
61	1.3	241	2	D87547	3-oxoadipate CoA-t
62	1.3	242	2	AE2314	ATP-binding protei
63	1.3	249	2	A99667	hypothetical prote
64	1.3	251	2	T47109	3-oxoadipate CoA-t
65	1.3	251	2	F72114	conserved hypothet
66	1.3	251	2	A86508	ACR family import
67	1.3	260	2	T35021	probable 3-oxoadip
68	1.3	279	2	E64109	dimethylsulfoxide
69	1.3	282	2	E95128	conserved hypothet
70	1.3	282	2	E97999	conserved hypothet
71	1.3	284	2	C70165	inositol monophosp
72	1.3	285	2	S08491	hypothetical prote
73	1.3	285	2	S74771	hypothetical prote
74	1.3	288	2	D85072	hypothetical prote
75	1.3	298	2	E97032	probable permease
76	1.3	299	2	E85517	probable LysR-like
77	1.3	300	2	AE0762	conserved hypothet
78	1.3	301	2	C75447	conserved hypothet
79	1.3	304	2	H75081	phosphate abc tran
80	1.3	305	1	H75285	probable phosphoes
81	1.3	307	2	E91206	probable ARAC-type
82	1.3	307	2	G86052	probable ARAC-type
83	1.3	307	2	A65170	hypothetical prote
84	1.3	307	2	AH3112	hypothetical prote
85	1.3	307	2	E98174	oligopeptide ABC t
86	1.3	313	2	B71390	NADH2 dehydrogenas
87	1.3	315	2	E64426	phosphate transpor
88	1.3	317	2	T11337	NADH2 dehydrogenas
89	1.3	318	1	QXBOIM	NADH2 dehydrogenas
90	1.3	318	2	A58888	NADH2 dehydrogenas
91	1.3	318	2	S47870	NADH2 dehydrogenas
92	1.3	318	2	T11428	NADH2 dehydrogenas
93	1.3	318	2	T11441	NADH2 dehydrogenas
94	1.3	318	2	T11440	NADH2 dehydrogenas
95	1.3	318	2	T11493	NADH2 dehydrogenas
96	1.3	318	2	T11402	NADH2 dehydrogenas
97	1.3	318	2	S41835	NADH2 dehydrogenas
98	1.3	318	2	S26151	NADH2 dehydrogenas
99	1.3	318	2	A58850	NADH2 dehydrogenas
100	1.3	318	2	S41820	NADH2 dehydrogenas
101	1.3	318	2	T45550	NADH2 dehydrogenas
102	1.3	318	2	T11363	NADH2 dehydrogenas

103	7	1.3	318	2	T11857	NADH2 dehydrogenas	176	7	1.3	528	2	G90569	hypothetical prote
104	7	1.3	318	2	T11247	NADH2 dehydrogenas	177	7	1.3	533	2	S10841	gene ND5 intron 1
105	7	1.3	318	2	T10972	NADH2 dehydrogenas	178	7	1.3	535	2	T03260	cytochrome P450 -
106	7	1.3	318	2	T11050	NADH2 dehydrogenas	179	7	1.3	535	2	T03246	cytochrome P450 -
107	7	1.3	321	2	T13811	NADH2 dehydrogenas	180	7	1.3	535	2	JC7858	GroEL protein - Te
108	7	1.3	321	2	S55004	NADH2 dehydrogenas	181	7	1.3	540	2	H98086	chaperonin GroEL [
109	7	1.3	321	2	AS0775	probable sugar kin	182	7	1.3	540	2	G95222	chaperonin, 60 kDa
110	7	1.3	322	2	S68128	NADH2 dehydrogenas	183	7	1.3	541	2	S72614	chaperonin 60 - Th
111	7	1.3	323	1	QXKL1M	NADH2 dehydrogenas	184	7	1.3	542	1	A70220	phosphotransferase
112	7	1.3	323	2	T11793	NADH2 dehydrogenas	185	7	1.3	544	2	JC6063	chaperonin groEL -
113	7	1.3	323	2	A34284	NADH2 dehydrogenas	186	7	1.3	544	2	B41884	58K heat shock pro
114	7	1.3	323	2	AS8892	NADH2 dehydrogenas	187	7	1.3	544	2	JC5130	heat shock protein
115	7	1.3	323	2	T09857	NADH2 dehydrogenas	188	7	1.3	544	2	B83720	class I heat-shock
116	7	1.3	323	2	T11820	NADH2 dehydrogenas	189	7	1.3	545	2	S10842	gene ND5 intron 2
117	7	1.3	323	2	T12486	hypothetical prote	190	7	1.3	558	2	G87450	conserved hypotet
118	7	1.3	324	2	S36002	NADH2 dehydrogenas	191	7	1.3	559	2	S62503	inorganic phosphat
119	7	1.3	324	2	S35462	NADH2 dehydrogenas	192	7	1.3	561	2	T49979	cytochrome P450-li
120	7	1.3	324	2	T09947	NADH2 dehydrogenas	193	7	1.3	563	2	T43850	probable sodium-de
121	7	1.3	324	2	C41608	hypothetical prote	194	7	1.3	568	2	AG2264	urease alpha chain
122	7	1.3	325	2	H86670	hypothetical prote	195	7	1.3	570	2	T49181	cyclophylin-like p
123	7	1.3	326	2	H71884	iron (III) dicitrat	196	7	1.3	576	2	H88548	protein ZK512.6 [i
124	7	1.3	326	2	A64631	iron(III) dicitrat	197	7	1.3	580	2	T39122	amino-acid permeas
125	7	1.3	334	2	T46885	3-methyl-2-oxobuta	198	7	1.3	594	2	T38114	pyruvate decarboxy
126	7	1.3	334	2	T36511	probable branched	199	7	1.3	600	2	E48951	nisin transport pr
127	7	1.3	336	2	C86921	conserved hypotet	200	7	1.3	605	2	T43191	probable pyruvate
128	7	1.3	343	2	T12425	NADH2 dehydrogenas	201	7	1.3	663	2	T38155	78 kd glucose regu
129	7	1.3	353	2	G02729	thrombopoietin - h	202	7	1.3	663	2	S20877	dnak-type molecula
130	7	1.3	353	2	I80105	thrombopoietin pre	203	7	1.3	674	2	I55476	growth potentiatin
131	7	1.3	359	2	T39944	hypothetical trans	204	7	1.3	678	2	B48089	growth arrest-spec
132	7	1.3	365	2	S50610	hypothetical prote	205	7	1.3	687	2	S74621	carbon dioxide con
133	7	1.3	370	2	AB0602	probable membrane	206	7	1.3	706	2	G90696	hypothetical prote
134	7	1.3	380	2	T29248	hypothetical prote	207	7	1.3	715	2	S10843	NADH2 dehydrogenas
135	7	1.3	384	2	A70805	probable lprN prot	208	7	1.3	720	2	C85547	probable cytoplasm
136	7	1.3	390	2	T50037	capsular polysacch	209	7	1.3	734	2	H86340	sugar transporter
137	7	1.3	392	2	JC7633	aryl hydrocarbon n	210	7	1.3	734	2	T51139	sugar transport pr
138	7	1.3	405	2	E89977	conserved hypotet	211	7	1.3	752	2	G02273	liv-1 protein - hu
139	7	1.3	407	2	F70696	hypothetical prote	212	7	1.3	753	2	T46614	chemotaxis protein
140	7	1.3	422	2	C82912	phosphoglycerate k	213	7	1.3	765	2	A69440	conserved hypotet
141	7	1.3	422	2	S73667	adhesin p1 precurs	214	7	1.3	778	2	T17679	proline-rich prote
142	7	1.3	423	2	T04915	CDP-diacylglycerol	215	7	1.3	781	2	T02272	hypothetical prote
143	7	1.3	423	2	G85255	CDP-diacylglycerol	216	7	1.3	792	2	S32244	X-Pro dipeptidyl-p
144	7	1.3	424	2	T07366	probable phosphati	217	7	1.3	796	2	T20393	hypothetical prote
145	7	1.3	427	2	T01455	CDP-diacylglycerol	218	7	1.3	803	2	B84931	DNA topoisomerase
146	7	1.3	428	2	S03767	cellulase (EC 3.2.	219	7	1.3	837	2	T12514	hypothetical prote
147	7	1.3	438	2	S73379	adhesin p1 precurs	220	7	1.3	847	2	G75270	cation-transportin
148	7	1.3	438	2	T04800	CDP-diacylglycerol	221	7	1.3	877	2	S65057	alpha-glucosidase
149	7	1.3	438	2	S73952	adhesin p1 precurs	222	7	1.3	924	2	S75284	chemotaxis protein
150	7	1.3	443	2	AD3050	MFS permease [alph	223	7	1.3	925	2	B83529	sensor/response re
151	7	1.3	443	2	S77166	hypothetical prote	224	7	1.3	1005	1	P1VXPJ	RNA 1 protein - pe
152	7	1.3	451	2	G69392	multidrug resistan	225	7	1.3	1085	2	F96712	hypothetical prote
153	7	1.3	458	2	T49004	transporter-like p	226	7	1.3	1144	2	T20218	hypothetical prote
154	7	1.3	459	2	R84935	argininosuccinate	227	7	1.3	1189	2	T51491	hypothetical prote
155	7	1.3	460	2	AB0968	sodium/galactoside	228	7	1.3	1275	2	B28096	line-1 protein ORF
156	7	1.3	461	2	D84971	phosphotransferase	229	7	1.3	1275	2	I38588	reverse transcript
157	7	1.3	461	2	G85059	probable sugar tra	230	7	1.3	1275	2	S65824	reverse transcript
158	7	1.3	462	2	B32840	anthranilate synth	231	7	1.3	1323	2	T18214	ATP binding casses
159	7	1.3	472	2	S28286	hypothetical prote	232	7	1.3	1627	1	IJYMAP	adhesin p1 precurs
160	7	1.3	473	2	S20612	triacylglycerol li	233	7	1.3	1635	2	A41480	adhesin p1, group
161	7	1.3	477	2	H83588	probable MFS trans	234	7	1.3	1682	2	A45380	sodium channel pro
162	7	1.3	477	2	H98235	probable mfs dicar	235	7	1.3	1683	2	T30885	complement compone
163	7	1.3	479	2	E91195	probable permease	236	7	1.3	1969	2	T08875	histidine kinase h
164	7	1.3	479	2	F86042	probable permease	237	7	1.3	2130	2	AB0821	probable exported
165	7	1.3	479	2	C65167	hypothetical 51.0	238	7	1.3	2145	2	S61041	glutamate synthase
166	7	1.3	487	2	G95315	probable [imported	239	7	1.3	2649	2	T51023	hypothetical prote
167	7	1.3	493	2	AE3077	aldehyde dehydroge	240	7	1.3	2890	2	B71846	dna-directed RNA p
168	7	1.3	493	2	D98209	aldhyde dehydroge	241	7	1.3	2890	2	F64669	DNA-directed RNA p
169	7	1.3	493	2	G88553	protein C38C10.2 [242	7	1.3	3122	2	T17202	DNA-directed DNA p
170	7	1.3	500	1	E69291	4-hydroxyphenylace	243	6	1.1	10	2	B60589	sperm-activating p
171	7	1.3	500	2	T49978	cytochrome P450-li	244	6	1.1	13	2	D61491	seed protein ws-4
172	7	1.3	504	2	T34411	hypothetical prote	245	6	1.1	18	2	A41877	LcrKc - Yersinia p
173	7	1.3	515	2	E71447	probable PSII D1 p	246	6	1.1	22	2	JP0069	ribosomal protein
174	7	1.3	517	2	A48250	U2AF-homologous pr	247	6	1.1	25	2	A60704	serine proteinase
175	7	1.3	523	2	S50479	26S proteasome reg	248	6	1.1	25	2	JP0067	ribosomal protein

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:55:49 ; Search time 115.414 Seconds

(without alignments)
1794.174 Million cell updates/sec

Title: US-09-776-865-2

Perfect score: 536

Sequence: 1 MAAGMTTPRPVQPARPGF.....LFAKEVQWALNDHGHRR 536

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1726216 seqs, 386330316 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications AA:*

```

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	536	100.0	536	9	US-09-776-865-2
2	536	100.0	536	17	US-10-823-506-8
3	495	92.4	495	9	US-09-915-181A-7
4	495	92.4	495	16	US-10-755-889-588
5	495	92.4	495	16	US-10-823-506-2
6	284	53.0	284	15	US-10-264-049-2477
7	217	40.5	272	15	US-10-264-237-2351
8	180	33.6	194	15	US-10-296-115-1200
9	61	11.4	495	9	US-09-776-865-4
10	61	11.4	495	17	US-10-823-506-4
11	14	2.6	495	17	US-10-823-506-10

12	14	2.6	495	17	US-10-823-506-12	Sequence 12, Appl
13	11	2.1	467	9	US-09-991-212A-3	Sequence 3, Appl
14	11	2.1	467	10	US-09-965-522-3	Sequence 3, Appl
15	11	2.1	467	17	US-10-877-818-3	Sequence 3, Appl
16	11	2.1	578	9	US-09-740-041-4	Sequence 4, Appl
17	11	2.1	578	14	US-10-389-967-4	Sequence 4, Appl
18	11	2.1	582	9	US-09-915-181A-4	Sequence 4, Appl
19	11	2.1	582	15	US-10-205-331-7	Sequence 7, Appl
20	11	2.1	582	16	US-10-734-731-10	Sequence 10, Appl
21	11	2.1	582	16	US-10-734-731-12	Sequence 10, Appl
22	11	2.1	582	16	US-10-734-731-14	Sequence 14, Appl
23	11	2.1	582	16	US-10-807-500-10	Sequence 10, Appl
24	11	2.1	582	16	US-10-807-500-12	Sequence 12, Appl
25	11	2.1	582	16	US-10-807-500-14	Sequence 14, Appl
26	10	1.9	544	15	US-10-369-493-5873	Sequence 5873, Ap
27	10	1.9	573	15	US-10-369-493-5871	Sequence 5871, Ap
28	10	1.9	573	15	US-10-369-493-5872	Sequence 5872, Ap
29	9	1.7	465	9	US-09-915-181A-8	Sequence 8, Appl
30	8	1.5	71	16	US-10-425-115-288336	Sequence 288336,
31	8	1.5	88	16	US-10-425-115-302951	Sequence 302951,
32	8	1.5	114	16	US-10-437-963-128557	Sequence 128557,
33	8	1.5	143	16	US-10-425-115-335986	Sequence 335986,
34	8	1.5	154	16	US-10-767-701-52130	Sequence 52130, A
35	8	1.5	232	15	US-10-425-114-43907	Sequence 43907, A
36	8	1.5	235	15	US-10-282-122A-56379	Sequence 56379, A
37	8	1.5	264	15	US-10-108-260A-3281	Sequence 3281, Ap
38	8	1.5	373	16	US-10-482-706-253	Sequence 253, App
39	8	1.5	384	15	US-10-424-599-183068	Sequence 183068,
40	8	1.5	478	16	US-10-437-963-180400	Sequence 180400,
41	8	1.5	479	15	US-10-369-493-23364	Sequence 23364, A
42	8	1.5	512	16	US-10-425-115-335979	Sequence 335979,
43	8	1.5	550	16	US-10-739-930-6619	Sequence 6619, Ap
44	8	1.5	560	9	US-09-991-212A-4	Sequence 4, Appl
45	8	1.5	560	10	US-09-915-181A-5	Sequence 5, Appl
46	8	1.5	560	15	US-10-314-790-5	Sequence 4, Appl
47	8	1.5	560	15	US-10-734-731-2	Sequence 2, Appl
48	8	1.5	560	16	US-10-734-731-4	Sequence 4, Appl
49	8	1.5	560	16	US-10-734-731-6	Sequence 6, Appl
50	8	1.5	560	16	US-10-734-731-8	Sequence 8, Appl
51	8	1.5	560	16	US-10-807-500-2	Sequence 2, Appl
52	8	1.5	560	16	US-10-807-500-4	Sequence 4, Appl
53	8	1.5	560	16	US-10-807-500-6	Sequence 6, Appl
54	8	1.5	560	16	US-10-807-500-8	Sequence 8, Appl
55	8	1.5	560	16	US-10-807-500-10	Sequence 10, Appl
56	8	1.5	560	17	US-10-877-818-4	Sequence 4, Appl
57	8	1.5	566	15	US-10-287-226-300	Sequence 300, App
58	8	1.5	566	16	US-10-437-963-112089	Sequence 112089,
59	8	1.5	589	9	US-09-740-041-2	Sequence 2, Appl
60	8	1.5	589	14	US-10-389-967-2	Sequence 2, Appl
61	8	1.5	589	18	US-10-499-731-28	Sequence 28, Appl
62	8	1.5	601	18	US-10-499-731-46	Sequence 46, Appl
63	8	1.5	850	9	US-09-915-181A-3	Sequence 3, Appl
64	8	1.5	940	16	US-10-437-963-155833	Sequence 155833,
65	8	1.5	949	17	US-10-732-923-18652	Sequence 18652, A
66	8	1.5	950	16	US-10-437-963-107881	Sequence 107881,
67	8	1.5	1031	16	US-10-437-963-107870	Sequence 107870,
68	8	1.5	4128	14	US-10-205-194-1	Sequence 1, Appl
69	7	1.3	13	15	US-10-469-101-69	Sequence 69, Appl
70	7	1.3	13	15	US-10-469-101-70	Sequence 70, Appl
71	7	1.3	13	15	US-10-469-101-71	Sequence 71, Appl
72	7	1.3	19	14	US-10-191-879-31	Sequence 31, Appl
73	7	1.3	33	10	US-09-983-802-663	Sequence 663, App
74	7	1.3	33	10	US-09-984-490-663	Sequence 663, App
75	7	1.3	33	11	US-09-973-278-444	Sequence 444, App
76	7	1.3	42	15	US-10-424-599-212978	Sequence 212978,
77	7	1.3	44	10	US-09-895-298-126	Sequence 126, App
78	7	1.3	44	16	US-10-885-039-126	Sequence 326, App
79	7	1.3	58	16	US-10-425-115-350119	Sequence 350119,
80	7	1.3	59	16	US-10-437-963-204921	Sequence 204921,
81	7	1.3	61	15	US-10-424-599-218670	Sequence 218670,
82	7	1.3	64	15	US-10-424-599-180080	Sequence 180080,
83	7	1.3	65	15	US-10-424-599-149839	Sequence 149839,
84	7	1.3	66	16	US-10-437-963-147591	Sequence 147591,

85	7	1.3	71	16	US-10-425-115-254514	Sequence 254514,	158	7	1.3	204	14	US-10-191-879-6	Sequence 6, Appli
86	7	1.3	71	16	US-10-425-115-318386	Sequence 318386,	159	7	1.3	204	15	US-10-104-047-2253	Sequence 2253, Ap
87	7	1.3	72	16	US-10-437-963-181072	Sequence 181072,	160	7	1.3	206	9	US-09-925-298-538	Sequence 538, App
88	7	1.3	76	15	US-10-424-599-255605	Sequence 255605,	161	7	1.3	206	14	US-10-102-806-538	Sequence 538, App
89	7	1.3	77	15	US-10-424-599-168281	Sequence 168281,	162	7	1.3	206	16	US-10-767-701-32217	Sequence 32217, A
90	7	1.3	78	16	US-10-425-115-357514	Sequence 375514,	163	7	1.3	207	15	US-10-264-237-1787	Sequence 1787, Ap
91	7	1.3	79	10	US-09-764-891-4448	Sequence 4448, Ap	164	7	1.3	212	15	US-10-424-599-160649	Sequence 160649, A
92	7	1.3	82	9	US-09-864-761-35759	Sequence 35759, A	165	7	1.3	223	9	US-09-925-300-1346	Sequence 1346, Ap
93	7	1.3	85	16	US-10-437-963-188238	Sequence 188238,	166	7	1.3	223	15	US-10-432-290-38	Sequence 38, Appl
94	7	1.3	87	15	US-10-424-599-148365	Sequence 148365,	167	7	1.3	223	18	US-10-276-7798-1	Sequence 1, Appli
95	7	1.3	88	16	US-10-425-115-253274	Sequence 253274,	168	7	1.3	223	15	US-10-432-290-24	Sequence 24, Appl
96	7	1.3	91	16	US-10-425-115-310462	Sequence 310462,	169	7	1.3	228	15	US-10-437-963-103966	Sequence 103966,
97	7	1.3	96	15	US-10-424-599-243646	Sequence 243646,	170	7	1.3	232	16	US-10-424-599-198853	Sequence 198853,
98	7	1.3	97	16	US-10-437-963-110377	Sequence 110377,	171	7	1.3	237	15	US-10-425-114-69641	Sequence 69641, A
99	7	1.3	97	16	US-10-425-115-199192	Sequence 199192,	172	7	1.3	237	15	US-10-312-273-141	Sequence 141, App
100	7	1.3	98	16	US-10-437-963-161222	Sequence 161222,	173	7	1.3	251	15	US-10-289-762-156	Sequence 156, App
101	7	1.3	101	15	US-10-424-599-199724	Sequence 199724,	174	7	1.3	251	15	US-10-424-599-180645	Sequence 180645,
102	7	1.3	103	15	US-10-094-749-2113	Sequence 2113, Ap	175	7	1.3	252	15	US-10-437-963-185811	Sequence 185811,
103	7	1.3	103	15	US-10-424-599-143096	Sequence 143096,	176	7	1.3	255	16	US-10-425-115-327619	Sequence 327619,
104	7	1.3	103	16	US-10-767-701-61821	Sequence 61821, A	177	7	1.3	258	16	US-10-282-122A-62062	Sequence 62062, A
105	7	1.3	105	16	US-10-425-115-333987	Sequence 333987,	178	7	1.3	259	15	US-10-156-761-9239	Sequence 9239, Ap
106	7	1.3	106	14	US-10-083-357-972	Sequence 972, App	179	7	1.3	260	14	US-10-425-115-319762	Sequence 319762,
107	7	1.3	108	16	US-10-425-115-253034	Sequence 253034,	180	7	1.3	268	16	US-10-425-115-327619	Sequence 327619,
108	7	1.3	110	16	US-10-425-115-246172	Sequence 246172,	181	7	1.3	271	16	US-10-425-114-40533	Sequence 40533, A
109	7	1.3	114	16	US-10-437-963-204470	Sequence 204470,	182	7	1.3	275	15	US-10-474-776-309	Sequence 309, App
110	7	1.3	111	11	US-09-864-408A-1266	Sequence 1266, Ap	183	7	1.3	282	16	US-10-472-928-2212	Sequence 2212, Ap
111	7	1.3	116	16	US-10-437-963-108779	Sequence 108779,	184	7	1.3	282	16	US-10-425-115-203803	Sequence 203803,
112	7	1.3	116	16	US-10-437-963-195396	Sequence 195396,	185	7	1.3	286	16	US-10-617-320-4120	Sequence 4120, Ap
113	7	1.3	117	16	US-10-425-115-365789	Sequence 365789,	186	7	1.3	287	18	US-10-322-281-494	Sequence 494, App
114	7	1.3	122	14	US-10-312-187-12	Sequence 12, Appl	187	7	1.3	292	16	US-09-813-153-128	Sequence 128, App
115	7	1.3	125	16	US-10-425-115-332340	Sequence 332340,	188	7	1.3	295	10	US-10-659-004-104	Sequence 104, App
116	7	1.3	126	15	US-10-282-122A-77481	Sequence 77481, A	189	7	1.3	302	17	US-10-369-386-5	Sequence 5, Appli
117	7	1.3	127	15	US-10-424-599-160647	Sequence 160647,	190	7	1.3	303	15	US-10-390-191-2	Sequence 2, Appli
118	7	1.3	129	16	US-10-767-701-39962	Sequence 39962, A	191	7	1.3	303	15	US-10-723-860-4548	Sequence 4548, Ap
119	7	1.3	133	16	US-10-767-701-49108	Sequence 49108, A	192	7	1.3	304	15	US-10-108-260A-4214	Sequence 4214, Ap
120	7	1.3	135	10	US-09-764-891-4759	Sequence 4759, Ap	193	7	1.3	306	15	US-10-304-928-16	Sequence 16, Appl
121	7	1.3	142	16	US-10-437-963-174382	Sequence 174382,	194	7	1.3	307	14	US-10-774-355A-2251	Sequence 2251, Ap
122	7	1.3	143	15	US-10-424-599-213589	Sequence 213589,	195	7	1.3	312	17	US-10-425-115-332328	Sequence 332328,
123	7	1.3	145	16	US-10-437-963-190613	Sequence 190613,	196	7	1.3	320	16	US-10-774-355A-2250	Sequence 2250, Ap
124	7	1.3	145	16	US-10-425-115-301015	Sequence 301015,	197	7	1.3	321	18	US-10-617-320-4282	Sequence 4282, Ap
125	7	1.3	146	16	US-10-437-963-137026	Sequence 137026,	198	7	1.3	323	15	US-10-369-493-21705	Sequence 21705, A
126	7	1.3	151	16	US-10-767-701-33254	Sequence 33254, A	199	7	1.3	323	15	US-10-282-122A-50442	Sequence 50442, A
127	7	1.3	152	15	US-10-425-115-206281	Sequence 206281,	200	7	1.3	326	9	US-09-815-242-11380	Sequence 11380, A
128	7	1.3	152	16	US-10-094-749-1733	Sequence 1733, Ap	201	7	1.3	326	9	US-09-815-242-11545	Sequence 11545, A
129	7	1.3	153	16	US-10-437-963-174385	Sequence 174385,	202	7	1.3	326	10	US-09-882-227-156	Sequence 156, App
130	7	1.3	153	16	US-10-425-115-301017	Sequence 301017,	203	7	1.3	326	15	US-10-282-122A-58828	Sequence 58828, A
131	7	1.3	155	9	US-09-925-299-831	Sequence 831, App	204	7	1.3	327	17	US-10-424-599-164057	Sequence 164057,
132	7	1.3	155	10	US-09-925-299-831	Sequence 831, App	205	7	1.3	327	17	US-09-925-300-1209	Sequence 1209, Ap
133	7	1.3	155	10	US-09-895-298-127	Sequence 127, App	206	7	1.3	332	14	US-10-774-355A-2059	Sequence 2059, Ap
134	7	1.3	155	14	US-10-106-698-5530	Sequence 5530, Ap	207	7	1.3	332	14	US-10-400-377-7	Sequence 7, Appli
135	7	1.3	155	15	US-10-424-599-279690	Sequence 279690,	208	7	1.3	332	14	US-10-400-708-7	Sequence 7, Appli
136	7	1.3	155	15	US-10-885-039-127	Sequence 127, App	209	7	1.3	332	14	US-10-298-148-7	Sequence 7, Appli
137	7	1.3	158	15	US-10-282-122A-63701	Sequence 63701, A	210	7	1.3	332	14	US-10-441-729-30	Sequence 30, Appl
138	7	1.3	159	16	US-10-437-963-129515	Sequence 129515,	211	7	1.3	332	15	US-10-469-101-1	Sequence 1, Appli
139	7	1.3	163	16	US-10-638-995-1	Sequence 1, Appli	212	7	1.3	332	15	US-10-638-995-2	Sequence 2, Appli
140	7	1.3	163	16	US-10-425-115-209403	Sequence 209403,	213	7	1.3	332	16	US-10-638-995-4	Sequence 4, Appli
141	7	1.3	163	17	US-10-811-492-1	Sequence 1, Appli	214	7	1.3	332	16	US-10-638-995-5	Sequence 5, Appli
142	7	1.3	169	15	US-10-425-114-72494	Sequence 72494, A	215	7	1.3	332	16	US-10-638-995-6	Sequence 6, Appli
143	7	1.3	170	15	US-10-425-114-41569	Sequence 41569, A	216	7	1.3	332	16	US-10-638-995-7	Sequence 7, Appli
144	7	1.3	171	15	US-10-424-599-143940	Sequence 143940,	217	7	1.3	332	16	US-10-638-995-8	Sequence 8, Appli
145	7	1.3	171	16	US-10-425-115-209425	Sequence 209425,	218	7	1.3	332	16	US-10-638-995-9	Sequence 9, Appli
146	7	1.3	173	14	US-10-156-761-11037	Sequence 11037, A	219	7	1.3	332	16	US-10-638-995-10	Sequence 10, Appl
147	7	1.3	174	16	US-10-767-701-34655	Sequence 34655, A	220	7	1.3	332	16	US-10-638-995-11	Sequence 11, Appl
148	7	1.3	175	15	US-10-424-599-214742	Sequence 214742,	221	7	1.3	332	16	US-10-638-995-12	Sequence 12, Appl
149	7	1.3	185	14	US-10-151-879-21	Sequence 21, Appl	222	7	1.3	332	16	US-10-638-995-13	Sequence 13, Appl
150	7	1.3	188	15	US-10-424-599-160648	Sequence 160648,	223	7	1.3	332	16	US-10-638-995-14	Sequence 14, Appl
151	7	1.3	194	16	US-10-437-963-174369	Sequence 174369,	224	7	1.3	332	16	US-10-638-995-15	Sequence 15, Appl
152	7	1.3	195	15	US-10-289-762-444	Sequence 444, App	225	7	1.3	332	16	US-10-638-995-16	Sequence 16, Appl
153	7	1.3	198	16	US-10-767-701-39961	Sequence 39961, A	226	7	1.3	332	16	US-10-638-995-17	Sequence 17, Appl
154	7	1.3	199	16	US-10-767-701-36019	Sequence 36019, A	227	7	1.3	332	16	US-10-638-995-18	Sequence 18, Appl
155	7	1.3	200	14	US-10-017-161-1078	Sequence 1078, Ap	228	7	1.3	332	16	US-10-638-995-19	Sequence 19, Appl
156	7	1.3	200	15	US-10-292-798-916	Sequence 916, App	229	7	1.3	332	16		
157	7	1.3	200	15	US-10-424-599-166917	Sequence 166917,	230	7	1.3	332	16		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:46:03 ; Search time 31.193 Seconds
(without alignments)
1282.720 Million cell updates/sec

Title: US-09-776-865-2

Perfect score: 536

Sequence: 1 MAGAMTPRPVQPARPGF.....LPKGEVQNALNDHGHRRH 536

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	536	100.0	536	US-09-359-167-8	Sequence 8, Appli
2	529	98.7	585	US-09-949-016-7705	Sequence 7705, Ap
3	495	92.4	495	US-09-359-167-2	Sequence 2, Appli
4	495	92.4	495	US-09-915-181A-7	Sequence 7, Appli
5	61	11.4	495	US-09-359-167-4	Sequence 4, Appli
6	14	2.6	495	US-09-359-167-10	Sequence 10, Appl
7	14	2.6	495	US-09-359-167-12	Sequence 12, Appl
8	11	2.1	467	US-08-805-118-3	Sequence 3, Appli
9	11	2.1	467	US-09-391-958-3	Sequence 3, Appli
10	11	2.1	480	US-08-724-394A-9	Sequence 9, Appli
11	11	2.1	578	US-09-740-041-4	Sequence 4, Appli
12	11	2.1	582	US-09-915-181A-4	Sequence 4, Appli
13	10	1.9	436	US-09-949-016-11448	Sequence 11448, A
14	10	1.9	470	US-08-724-394A-10	Sequence 10, Appl
15	9	1.7	465	US-09-915-181A-8	Sequence 8, Appli
16	8	1.5	140	US-09-270-767-33069	Sequence 33069, A
17	8	1.5	140	US-09-270-767-48286	Sequence 48286, A
18	8	1.5	173	US-09-902-540-10210	Sequence 10210, A
19	8	1.5	218	US-09-902-540-11754	Sequence 11754, A
20	8	1.5	318	US-09-902-540-11509	Sequence 11509, A
21	8	1.5	380	US-09-949-016-7053	Sequence 7053, Ap
22	8	1.5	464	US-09-107-532A-5509	Sequence 5509, Ap
23	8	1.5	497	US-09-949-016-6616	Sequence 6616, Ap
24	8	1.5	501	US-09-252-991A-21596	Sequence 21596, A
25	8	1.5	560	US-08-647-484-2	Sequence 2, Appli
26	8	1.5	560	US-08-647-481-2	Sequence 2, Appli
27	8	1.5	560	US-08-430-033A-2	Sequence 2, Appli

28	8	1.5	560	2	US-08-805-118-4	Sequence 4, Appli
29	8	1.5	560	3	US-09-391-958-4	Sequence 4, Appli
30	8	1.5	560	4	US-09-915-181A-5	Sequence 5, Appli
31	8	1.5	560	5	PCT-US96-05792-2	Sequence 2, Appli
32	8	1.5	567	4	US-09-949-016-11354	Sequence 11354, A
33	8	1.5	589	4	US-09-740-041-2	Sequence 2, Appli
34	8	1.5	687	4	US-09-252-991A-26187	Sequence 26187, A
35	8	1.5	850	4	US-09-915-181A-3	Sequence 3, Appli
36	7	1.3	33	3	US-09-227-357-663	Sequence 663, App
37	7	1.3	48	4	US-09-270-767-60724	Sequence 60724, A
38	7	1.3	60	4	US-09-621-976-6063	Sequence 6063, Ap
39	7	1.3	102	4	US-09-621-976-7063	Sequence 7063, Ap
40	7	1.3	102	4	US-09-710-279-2592	Sequence 2592, Ap
41	7	1.3	110	4	US-09-248-796A-14320	Sequence 14320, A
42	7	1.3	117	4	US-09-270-767-39798	Sequence 39798, A
43	7	1.3	117	4	US-09-270-767-55015	Sequence 55015, A
44	7	1.3	122	3	US-09-134-001C-4987	Sequence 4987, Ap
45	7	1.3	122	4	US-09-949-016-10642	Sequence 10642, A
46	7	1.3	127	4	US-09-489-039A-13540	Sequence 13540, A
47	7	1.3	136	4	US-09-513-999C-6210	Sequence 6210, Ap
48	7	1.3	157	4	US-09-270-767-44201	Sequence 44201, A
49	7	1.3	157	4	US-09-270-767-61037	Sequence 61037, A
50	7	1.3	165	1	US-08-413-803-29	Sequence 29, Appl
51	7	1.3	174	5	PCT-US95-03776-27	Sequence 27, Appl
52	7	1.3	187	3	US-09-194-185-1	Sequence 1, Appli
53	7	1.3	195	1	US-08-388-779A-4	Sequence 4, Appli
54	7	1.3	195	1	US-08-591-070A-4	Sequence 4, Appli
55	7	1.3	195	2	US-08-927-855-4	Sequence 4, Appli
56	7	1.3	195	4	US-09-198-452A-444	Sequence 444, App
57	7	1.3	199	4	US-09-248-796A-16632	Sequence 16632, A
58	7	1.3	200	4	US-09-438-185A-427	Sequence 427, App
59	7	1.3	204	4	US-09-549-831-6	Sequence 6, Appli
60	7	1.3	218	4	US-09-134-000C-6768	Sequence 6768, Ap
61	7	1.3	225	4	US-09-270-767-45231	Sequence 45231, A
62	7	1.3	228	4	US-09-540-236-2798	Sequence 2798, Ap
63	7	1.3	232	4	US-09-252-991A-23718	Sequence 23718, A
64	7	1.3	234	4	US-09-270-767-45527	Sequence 45527, A
65	7	1.3	251	4	US-09-198-452A-156	Sequence 156, App
66	7	1.3	251	4	US-09-540-236-2306	Sequence 2306, Ap
67	7	1.3	252	4	US-09-252-991A-18642	Sequence 18642, A
68	7	1.3	254	4	US-09-438-185A-139	Sequence 139, App
69	7	1.3	260	4	US-09-252-991A-32744	Sequence 32744, A
70	7	1.3	265	4	US-09-489-039A-11328	Sequence 11328, A
71	7	1.3	279	3	US-08-549-515-7	Sequence 7, Appli
72	7	1.3	282	4	US-09-583-110-3800	Sequence 3800, Ap
73	7	1.3	287	4	US-09-107-433-4120	Sequence 4120, Ap
74	7	1.3	291	4	US-09-252-991A-28572	Sequence 28572, A
75	7	1.3	301	4	US-09-902-540-15018	Sequence 15018, A
76	7	1.3	302	4	US-09-248-796A-16767	Sequence 16767, A
77	7	1.3	303	1	US-08-467-607-3	Sequence 3, Appli
78	7	1.3	303	2	US-08-469-362-3	Sequence 3, Appli
79	7	1.3	303	2	US-08-850-392-3	Sequence 3, Appli
80	7	1.3	307	4	US-09-522-714-16	Sequence 16, Appl
81	7	1.3	307	4	US-09-489-039A-10116	Sequence 10116, A
82	7	1.3	310	4	US-09-248-796A-17322	Sequence 17322, A
83	7	1.3	312	3	US-08-875-533-67	Sequence 67, Appl
84	7	1.3	312	4	US-09-252-991A-22205	Sequence 22205, A
85	7	1.3	321	4	US-09-107-433-4282	Sequence 4282, Ap
86	7	1.3	328	3	US-08-875-533-37	Sequence 37, Appl
87	7	1.3	332	3	US-08-875-533-1	Sequence 1, Appli
88	7	1.3	332	3	US-08-875-533-2	Sequence 2, Appli
89	7	1.3	332	3	US-08-875-533-3	Sequence 3, Appli
90	7	1.3	332	3	US-08-875-533-4	Sequence 4, Appli
91	7	1.3	332	3	US-08-875-533-5	Sequence 5, Appli
92	7	1.3	332	3	US-08-875-533-6	Sequence 6, Appli
93	7	1.3	332	3	US-08-875-533-7	Sequence 7, Appli
94	7	1.3	332	3	US-08-875-533-8	Sequence 8, Appli
95	7	1.3	332	3	US-08-875-533-9	Sequence 9, Appli
96	7	1.3	332	3	US-08-875-533-10	Sequence 10, Appl
97	7	1.3	332	3	US-08-875-533-11	Sequence 11, Appl
98	7	1.3	332	3	US-08-875-533-22	Sequence 22, Appl
99	7	1.3	332	3	US-08-875-533-66	Sequence 66, Appl
100	7	1.3	332	4	US-09-462-941-7	Sequence 7, Appli

101	7	1.3	332	4	US-09-510-238A-3	Sequence 3, Appli	174	7	1.3	947	4	US-09-252-991A-29359	Sequence 29359, A
102	7	1.3	332	5	PCT-US95-03776-25	Sequence 25, Appli	175	7	1.3	1055	4	US-09-404-967C-2	Sequence 2, Appli
103	7	1.3	334	1	US-08-482-385A-7	Sequence 7, Appli	176	7	1.3	1075	4	US-09-404-967C-4	Sequence 4, Appli
104	7	1.3	335	4	US-09-134-000C-3742	Sequence 3742, Ap	177	7	1.3	1121	4	US-09-949-016-11048	Sequence 11048, A
105	7	1.3	344	4	US-09-949-016-9639	Sequence 9639, Ap	178	7	1.3	1169	1	US-08-315-468-4	Sequence 4, Appli
106	7	1.3	348	4	US-09-252-991A-26643	Sequence 26643, A	179	7	1.3	1255	4	US-09-902-540-12628	Sequence 12628, A
107	7	1.3	351	2	US-08-843-530B-31	Sequence 31, Appl	180	7	1.3	1374	4	US-09-900-425B-2	Sequence 2, Appli
108	7	1.3	351	4	US-09-636-728-28	Sequence 28, Appl	181	7	1.3	1421	3	US-09-335-409-2	Sequence 2, Appli
109	7	1.3	353	1	US-08-330-517-2	Sequence 2, Appli	182	7	1.3	1421	3	US-09-568-102-2	Sequence 2, Appli
110	7	1.3	353	1	US-08-347-029-4	Sequence 4, Appli	183	7	1.3	1421	3	US-09-567-969-2	Sequence 2, Appli
111	7	1.3	353	1	US-08-388-779A-2	Sequence 2, Appli	184	7	1.3	1421	3	US-09-568-480-2	Sequence 2, Appli
112	7	1.3	353	1	US-08-484-246-1	Sequence 1, Appli	185	7	1.3	1421	3	US-09-568-486-2	Sequence 2, Appli
113	7	1.3	353	1	US-08-491-070A-2	Sequence 2, Appli	186	7	1.3	1421	3	US-09-568-472-2	Sequence 2, Appli
114	7	1.3	353	1	US-08-413-803-25	Sequence 25, Appl	187	7	1.3	1421	3	US-09-567-899-2	Sequence 2, Appli
115	7	1.3	353	1	US-08-321-488A-25	Sequence 25, Appl	188	7	1.3	1627	1	US-07-665-792E-9	Sequence 9, Appli
116	7	1.3	353	2	US-08-414-161B-2	Sequence 2, Appli	189	7	1.3	1865	1	US-08-588-985-2	Sequence 2, Appli
117	7	1.3	353	2	US-08-457-254-4	Sequence 4, Appli	190	7	1.3	1865	1	US-08-971-988-2	Sequence 2, Appli
118	7	1.3	353	2	US-08-457-254-6	Sequence 6, Appli	191	7	1.3	1865	4	US-09-949-016-6069	Sequence 6069, Ap
119	7	1.3	353	2	US-08-927-855-2	Sequence 2, Appli	192	6	1.1	9	4	US-10-141-645-84	Sequence 84, Appl
120	7	1.3	353	3	US-08-482-212-2	Sequence 2, Appli	193	6	1.1	10	4	US-09-653-465B-11	Sequence 11, Appl
121	7	1.3	353	3	US-08-461-819-6	Sequence 6, Appli	194	6	1.1	10	4	US-09-191-593-41	Sequence 41, Appl
122	7	1.3	353	3	US-08-471-045-55	Sequence 55, Appl	195	6	1.1	12	4	US-09-674-973A-282	Sequence 282, App
123	7	1.3	353	3	US-08-469-712A-55	Sequence 55, Appl	196	6	1.1	13	4	US-09-190-976B-2	Sequence 2, Appli
124	7	1.3	353	3	US-08-999-927-2	Sequence 2, Appli	197	6	1.1	15	3	US-09-124-541-5	Sequence 5, Appli
125	7	1.3	353	3	US-08-999-927-6	Sequence 6, Appli	198	6	1.1	15	4	US-09-663-326-5	Sequence 5, Appli
126	7	1.3	353	3	US-08-461-819-2	Sequence 2, Appli	199	6	1.1	16	2	US-08-642-406A-12	Sequence 12, Appl
127	7	1.3	353	3	US-08-446-871-55	Sequence 55, Appl	200	6	1.1	16	3	US-09-199-534-12	Sequence 12, Appl
128	7	1.3	353	3	US-08-468-910-55	Sequence 55, Appl	201	6	1.1	16	4	US-09-199-534-12	Sequence 12, Appl
129	7	1.3	353	4	US-08-761-907-55	Sequence 55, Appl	202	6	1.1	20	3	US-08-744-419-10	Sequence 10, Appl
130	7	1.3	353	4	US-08-872-702-13	Sequence 13, Appl	203	6	1.1	26	4	US-09-257-179-101	Sequence 101, App
131	7	1.3	353	5	PCT-US94-08806-19	Sequence 19, Appl	204	6	1.1	28	3	US-09-380-061B-8	Sequence 8, Appli
132	7	1.3	353	5	PCT-US94-08806-29	Sequence 29, Appl	205	6	1.1	28	4	US-09-091-097-53	Sequence 53, Appl
133	7	1.3	353	5	PCT-US95-01829-2	Sequence 2, Appli	206	6	1.1	34	1	US-07-745-206A-5	Sequence 5, Appli
134	7	1.3	353	5	PCT-US95-01829-6	Sequence 6, Appli	207	6	1.1	34	1	US-08-455-543A-50	Sequence 50, Appl
135	7	1.3	353	5	PCT-US95-14932-4	Sequence 4, Appli	208	6	1.1	34	2	US-08-223-305C-50	Sequence 50, Appl
136	7	1.3	353	5	PCT-US95-16626-4	Sequence 4, Appli	209	6	1.1	34	2	US-08-311-363-5	Sequence 5, Appli
137	7	1.3	353	5	PCT-US95-16626-6	Sequence 6, Appli	210	6	1.1	36	2	US-08-846-762-59	Sequence 59, Appl
138	7	1.3	362	4	US-09-248-796A-26034	Sequence 26034, A	211	6	1.1	40	4	US-09-602-777A-12	Sequence 12, Appl
139	7	1.3	367	4	US-09-252-991A-16628	Sequence 16628, A	212	6	1.1	41	4	US-09-270-767-36936	Sequence 36936, A
140	7	1.3	404	4	US-09-543-681A-7864	Sequence 7864, Ap	213	6	1.1	41	4	US-09-270-767-52153	Sequence 52153, A
141	7	1.3	431	1	US-08-311-023-2	Sequence 2, Appli	214	6	1.1	42	3	US-08-822-774-41	Sequence 41, Appl
142	7	1.3	437	1	US-09-252-991A-20861	Sequence 20861, A	215	6	1.1	42	3	US-09-632-711-41	Sequence 41, Appl
143	7	1.3	469	4	US-09-489-039A-10553	Sequence 10553, A	216	6	1.1	42	3	US-09-632-703B-41	Sequence 41, Appl
144	7	1.3	473	4	US-09-411-132A-8	Sequence 8, Appli	217	6	1.1	42	4	US-09-632-702-41	Sequence 41, Appl
145	7	1.3	512	4	US-09-902-540-15789	Sequence 15789, A	218	6	1.1	42	4	US-09-399-003-41	Sequence 41, Appl
146	7	1.3	523	4	US-09-538-092-234	Sequence 234, App	219	6	1.1	43	4	US-09-489-847-225	Sequence 225, App
147	7	1.3	540	4	US-09-252-991A-22555	Sequence 22555, A	220	6	1.1	44	4	US-09-205-258-257	Sequence 257, App
148	7	1.3	540	4	US-09-583-110-3685	Sequence 3685, Ap	221	6	1.1	50	3	US-09-282-856A-1	Sequence 1, Appli
149	7	1.3	552	4	US-09-270-767-45540	Sequence 45540, A	222	6	1.1	51	4	US-09-270-767-57646	Sequence 57646, A
150	7	1.3	563	4	US-09-915-181A-6	Sequence 6, Appli	223	6	1.1	52	4	US-09-621-976-6359	Sequence 6359, Ap
151	7	1.3	574	4	US-09-248-796A-16849	Sequence 16849, A	224	6	1.1	54	4	US-09-270-767-57137	Sequence 57137, A
152	7	1.3	576	3	US-08-864-785-1	Sequence 1, Appli	225	6	1.1	56	3	US-09-177-249-166	Sequence 166, App
153	7	1.3	641	4	US-09-613-303-51	Sequence 51, Appl	226	6	1.1	56	4	US-09-755-665-53	Sequence 53, Appl
154	7	1.3	641	4	US-10-267-311-51	Sequence 51, Appl	227	6	1.1	56	4	US-09-812-283-166	Sequence 166, App
155	7	1.3	663	1	US-08-441-139-7	Sequence 7, Appli	228	6	1.1	59	4	US-09-270-767-41846	Sequence 41846, A
156	7	1.3	678	1	US-08-282-141-2	Sequence 2, Appli	229	6	1.1	59	4	US-09-270-767-57088	Sequence 57088, A
157	7	1.3	678	1	US-08-435-434-2	Sequence 2, Appli	230	6	1.1	61	1	US-08-095-898-6	Sequence 6, Appli
158	7	1.3	678	1	US-08-435-436-2	Sequence 2, Appli	231	6	1.1	61	4	US-09-270-767-60545	Sequence 60545, A
159	7	1.3	678	2	US-08-438-863-2	Sequence 2, Appli	232	6	1.1	61	5	PCT-US92-09439-6	Sequence 6, Appli
160	7	1.3	678	2	US-08-438-864-2	Sequence 2, Appli	233	6	1.1	62	4	US-09-661-322A-26	Sequence 26, Appl
161	7	1.3	678	3	US-08-438-862-2	Sequence 2, Appli	234	6	1.1	62	4	US-09-270-767-59377	Sequence 59377, A
162	7	1.3	678	3	US-08-628-747-2	Sequence 2, Appli	235	6	1.1	62	4	US-09-248-796A-25193	Sequence 25193, A
163	7	1.3	678	3	US-08-402-253-2	Sequence 2, Appli	236	6	1.1	64	4	US-09-661-322A-32	Sequence 32, Appl
164	7	1.3	678	3	US-08-443-866B-2	Sequence 2, Appli	237	6	1.1	64	4	US-09-248-796A-26590	Sequence 26590, A
165	7	1.3	693	4	US-09-252-991A-20348	Sequence 20348, A	238	6	1.1	66	4	US-09-489-039A-11065	Sequence 11065, A
166	7	1.3	697	4	US-09-252-991A-21106	Sequence 21106, A	239	6	1.1	66	4	US-09-673-395A-165	Sequence 165, App
167	7	1.3	722	4	US-09-252-991A-26839	Sequence 26839, A	240	6	1.1	66	4	US-09-513-999C-5262	Sequence 5262, Ap
168	7	1.3	755	4	US-09-642-034-5	Sequence 5, Appli	241	6	1.1	67	3	US-09-134-001C-4786	Sequence 4786, Ap
169	7	1.3	760	4	US-09-252-991A-27790	Sequence 27790, A	242	6	1.1	67	4	US-09-489-039A-10141	Sequence 10141, A
170	7	1.3	778	4	US-09-248-796A-16014	Sequence 16014, A	243	6	1.1	68	4	US-09-252-991A-32975	Sequence 32975, A
171	7	1.3	834	4	US-09-404-967C-5	Sequence 5, Appli	244	6	1.1	69	4	US-09-107-532A-6142	Sequence 6142, Ap
172	7	1.3	877	1	US-08-430-925A-6	Sequence 4, Appli	245	6	1.1	69	4	US-09-248-796A-14723	Sequence 14723, A
173	7	1.3	937	4	US-09-949-016-8286	Sequence 8286, Ap	246	6	1.1	70	4	US-09-621-976-6964	Sequence 6964, Ap

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:29:52 ; Search time 116.974 Seconds

(without alignments)
1772.222 Million cell updates/sec

Title: US-09-776-865-2

Perfect score: 536

Sequence: 1 MAAGAMTPRPVQPARPGGF.....LFAKGEVQWALNDHHGRH 536

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A Genesecp16Dec04:*

1: Genesecp1980s:*

2: Genesecp1990s:*

3: Genesecp2000s:*

4: Genesecp2001s:*

5: Genesecp2002s:*

6: Genesecp2003as:*

7: Genesecp2003bs:*

8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	536	100.0	536	3	AAY45089 Human GBS
2	536	100.0	536	4	AAY45089 Human GBS
3	495	92.4	495	3	AAY45087 Partial h
4	495	92.4	495	4	AAB66967 Human AGT
5	495	92.4	495	8	ADJ75516 Marker ge
6	495	92.4	495	8	ADRI4587 Human NF-
7	495	92.4	495	8	ADP25217 PRO polyp
8	394	73.5	495	4	AAM38959 Human pol
9	314	58.6	314	4	AAG65238 Human sod
10	304	56.7	309	4	AAM40745 Human pol
11	284	53.0	284	5	ABP41345 Human ova
12	273	50.9	277	4	AAM93914 Human pro
13	273	50.9	277	8	ADL32036 Human pro
14	217	40.5	272	5	ABB89975 Human pol
15	180	33.6	194	4	AAM25685 Human pro
16	144	26.9	343	4	ABG12999 Novel hum
17	144	26.9	343	6	ABO00820 Polypepti
18	134	25.0	256	4	ABG02032 Novel hum
19	124	23.1	124	4	AAG81307 Human AFP
20	61	11.4	495	3	AAY45088 Sheep GBS
21	61	11.4	495	4	AAG06519 Sheep gro
22	15	2.8	15	4	AAG65239 Human sod
23	14	2.6	495	3	AAY45090 Human/She
24	11	2.1	467	8	ADP12952 Protein e
25	11	2.1	578	7	ADG88331 Rat trans

26	11	2.1	582	4	AAM79273 Human pro
27	11	2.1	582	6	ABG74796 Rat DNPI
28	11	2.1	582	6	ABG74795 Human DNP
29	11	2.1	582	6	ABG74797 Murine DN
30	11	2.1	582	6	ABM04787 Rat Na-Ge
31	11	2.1	582	7	ADCI15494 Mouse DNP
32	11	2.1	582	7	ADCI15492 Rat DNPI
33	11	2.1	582	7	ADCI15490 Human DNP
34	11	2.1	582	7	ADCI15490 Human VGL
35	10	1.9	387	8	ADP29764 Human sec
36	10	1.9	436	2	AAM78919 Human hae
37	10	1.9	436	7	ADF90914 Human hep
38	10	1.9	544	8	ADN23220 Bacterial
39	10	1.9	573	8	ADN23219 Bacterial
40	10	1.9	573	8	ADN23218 Bacterial
41	9	1.7	462	4	ABE70143 Drosophill
42	8	1.5	91	4	ABG28472 Novel hum
43	8	1.5	235	6	ABU28455 Protein e
44	8	1.5	264	7	ADM04596 Human pro
45	8	1.5	373	7	ADB80215 Mycobacte
46	8	1.5	408	3	AAG39996 Arabidops
47	8	1.5	408	3	AAG14816 Arabidops
48	8	1.5	444	3	AAG39995 Arabidops
49	8	1.5	444	3	AAG14815 Arabidops
50	8	1.5	464	7	ADC95882 E faeciu
51	8	1.5	479	8	ADS44934 Bacterial
52	8	1.5	497	4	ABB70142 Drosophill
53	8	1.5	501	7	ABO72850 Pseudomon
54	8	1.5	502	4	ABB65873 Drosophill
55	8	1.5	502	4	ABB60525 Drosophill
56	8	1.5	550	3	AAG39994 Arabidops
57	8	1.5	550	3	AAG14814 Arabidops
58	8	1.5	559	4	ABB58701 Drosophill
59	8	1.5	560	2	AAM05148 Human bra
60	8	1.5	560	2	AAM70500 Human sod
61	8	1.5	560	6	ABG74794 Murine BN
62	8	1.5	560	6	ABG74793 Rat BNPI
63	8	1.5	560	6	ABG74792 Human BNP
64	8	1.5	560	6	ABG74791 Human BNP
65	8	1.5	560	7	ADC15488 Mouse BNP
66	8	1.5	560	7	ADC15484 Human BNP
67	8	1.5	560	7	ADC15482 Human BNP
68	8	1.5	560	7	ADC15486 Rat BNPI
69	8	1.5	560	7	ADD01474 Human VGL
70	8	1.5	566	7	ADJ95072 Novel NOV
71	8	1.5	567	4	AAO13870 Human pol
72	8	1.5	589	5	AAU99329 Human tra
73	8	1.5	589	5	AAO30994 Human tra
74	8	1.5	589	7	ADD01392 Human TCH
75	8	1.5	589	7	ADG88329 Human tra
76	8	1.5	589	8	ADR10021 Human pro
77	8	1.5	601	7	ADD01410 Mouse TCH
78	8	1.5	687	7	ABO77441 Pseudomon
79	8	1.5	860	5	ABB07689 Rat glura
80	8	1.5	953	4	ABB63934 Drosophill
81	8	1.5	2015	4	ABB65890 Drosophill
82	8	1.5	4128	7	ADB85120 Mouse DNA
83	7	1.3	13	6	ABP81116 Human TPO
84	7	1.3	13	6	ABP81117 Human TPO
85	7	1.3	13	6	ABP81115 Human TPO
86	7	1.3	18	3	ABY88326 Human Cat
87	7	1.3	19	7	ADC39960 Yellow gr
88	7	1.3	24	3	AAY88323 Human Cat
89	7	1.3	33	2	AAV03007 Fragment
90	7	1.3	33	2	ADA07842 Human sec
91	7	1.3	33	8	ADA41322 Novel hum
92	7	1.3	40	4	AAB64754 Gene 36 h
93	7	1.3	40	5	AAB89268 Insulin/i
94	7	1.3	40	6	ABJ25329 SAC isola
95	7	1.3	44	3	ABJ25329 Human sec
96	7	1.3	44	5	AAU090770 Insulin/i
97	7	1.3	52	2	AAH81401 Artificial
98	7	1.3	59	3	AAB38279 Gene 20 h

99	7	1.3	67	3	AA540211	Aab40211	Gene 17 h	172	7	1.3	163	8	ADM09216	Human TPO
100	7	1.3	76	4	AAU65452	Aau65452	Propionib	173	7	1.3	164	2	AAR81348	[H1533, T
101	7	1.3	76	6	ABM61971	Abm61971	Propionib	174	7	1.3	164	2	AAR81349	[H1533, A
102	7	1.3	78	4	AAU40931	Aau40931	Propionib	175	7	1.3	164	2	AAR81352	[GLY116,
103	7	1.3	78	6	ABM37450	Abm37450	Propionib	176	7	1.3	164	2	AAR81350	[H1533, G
104	7	1.3	79	4	AAW95790	Aaw95790	Human rep	177	7	1.3	164	2	AAR81353	[GLY116,
105	7	1.3	79	4	AB996321	Ab996321	Human tes	178	7	1.3	164	2	AAR81351	[GLY116,
106	7	1.3	81	5	ABP08020	Abp08020	Human ORF	179	7	1.3	165	2	AAW04868	Human thr
107	7	1.3	82	4	AAW16057	Aaw16057	Peptide #	180	7	1.3	165	2	AAW04868	Human thr
108	7	1.3	82	4	ABW35049	Abw35049	Peptide #	181	7	1.3	165	3	AAW55013	Human will
109	7	1.3	82	4	AAW28554	Aaw28554	Peptide #	182	7	1.3	165	7	ADC54795	Human thr
110	7	1.3	82	4	ABW29865	Abw29865	Peptide #	183	7	1.3	166	2	AAR97764	Human thr
111	7	1.3	82	4	ABW20461	Abw20461	Protein #	184	7	1.3	171	2	AAR97770	Human thr
112	7	1.3	82	4	AAW68235	Aaw68235	Human bon	185	7	1.3	171	2	AAR97765	Human thr
113	7	1.3	82	4	AAW55864	Aaw55864	Human bra	186	7	1.3	174	2	AAR80823	Human MGD
114	7	1.3	82	4	ABG49886	Abg49886	Human liv	187	7	1.3	175	2	AAW12953	[N30, T32,
115	7	1.3	82	4	AAW03784	Aaw03784	Peptide #	188	7	1.3	175	6	ABP78104	N. gonorr
116	7	1.3	82	5	ABG37771	Abg37771	Human pep	189	7	1.3	176	2	AAR97771	Human thr
117	7	1.3	86	4	ABW65993	Abw65993	Human sec	190	7	1.3	179	2	AAR81390	Human TPO
118	7	1.3	99	3	AAW22278	Aaw22278	Arabidops	191	7	1.3	179	5	ABU50929	Helicobac
119	7	1.3	100	4	ABG27184	Abg27184	Novel hum	192	7	1.3	180	2	AAR97766	Human thr
120	7	1.3	102	4	AAW82749	Aaw82749	S. epider	193	7	1.3	185	1	AAW90694	Sequence
121	7	1.3	103	4	AAW86466	Aaw86466	Human imm	194	7	1.3	185	2	AAR97772	Human thr
122	7	1.3	103	6	ADA54545	Ada54545	Human pro	195	7	1.3	185	7	ADC39950	Yellow gr
123	7	1.3	106	5	ADH32514	Adh32514	Yeast smo	196	7	1.3	186	1	AAW90616	Fish grow
124	7	1.3	107	5	ADK36539	Adk36539	Novel hum	197	7	1.3	187	1	AAW1244	Sequence
125	7	1.3	110	8	ADQ66018	Adq66018	Novel hum	198	7	1.3	187	2	AAW71372	Synthetic
126	7	1.3	114	2	AAW54390	Aaw54390	Actinomad	199	7	1.3	187	2	AAW49019	Korean ro
127	7	1.3	115	5	ABP31660	Abp31660	Human ORP	200	7	1.3	188	2	AAW07353	Fish grow
128	7	1.3	119	3	AAW22277	Aaw22277	Arabidops	201	7	1.3	188	2	AAW24137	Fish grow
129	7	1.3	119	3	AAW08053	Aaw08053	Arabidops	202	7	1.3	192	8	ABM80182	Tumour-as
130	7	1.3	122	5	AAU81759	Aau81759	Partial p	203	7	1.3	195	2	AAR99859	Truncated
131	7	1.3	122	5	ABP40142	Abp40142	Staphyloc	204	7	1.3	195	2	AAW06827	Human tru
132	7	1.3	122	8	ADSO7186	Adso7186	Staphyloc	205	7	1.3	195	2	AAW35026	Chlamydia
133	7	1.3	126	6	ABU49557	Abu49557	Protein e	206	7	1.3	200	1	AAW90732	Fish grow
134	7	1.3	127	7	ABO67023	Ab067023	Klebsiell	207	7	1.3	200	2	AAW12954	[N30, T32,
135	7	1.3	131	4	AAU20576	Aau20576	Human sec	208	7	1.3	200	4	AAW59504	Human sec
136	7	1.3	132	4	AAW59519	Aaw59519	Human sec	209	7	1.3	200	7	ADC86463	Human GPC
137	7	1.3	132	7	ADM25718	Adm25718	Hypether	210	7	1.3	202	4	ABB63463	Drosophil
138	7	1.3	133	4	ABG11406	Abg11406	Novel hum	211	7	1.3	203	1	AAW90617	Fish grow
139	7	1.3	135	4	AAW96101	Aaw96101	Human rep	212	7	1.3	203	2	AAW07248	Fish grow
140	7	1.3	136	3	AAW02129	Aaw02129	Human sec	213	7	1.3	203	3	AAU77282	Spotted f
141	7	1.3	140	2	AAW94926	Aaw94926	Human phe	214	7	1.3	204	1	AAW95638	Fish grow
142	7	1.3	150	2	AAW80822	Aaw80822	Human MGD	215	7	1.3	204	2	AAW10912	Fish grow
143	7	1.3	152	6	ADA54165	Ada54165	Human pro	216	7	1.3	204	2	AAW27340	Yellow ta
144	7	1.3	155	3	AAW53291	Aaw53291	Human col	217	7	1.3	204	2	AAW25120	Yellowtai
145	7	1.3	155	3	AAW24501	Aaw24501	Human sec	218	7	1.3	204	2	AAW49020	Korean ro
146	7	1.3	155	4	AAW74756	Aaw74756	Human col	219	7	1.3	204	5	ABG31737	Yellow gr
147	7	1.3	157	8	ADR10198	Adr10198	Human pro	220	7	1.3	204	7	ABG31734	Yellow gr
148	7	1.3	158	6	AAW97763	Aaw97763	Human thr	221	7	1.3	204	7	ADB64099	Human pro
149	7	1.3	158	6	ABU35777	Abu35777	Protein e	222	7	1.3	204	7	ADC50041	Orange-sp
150	7	1.3	162	2	AAW81346	Aaw81346	[Delta-Ar	223	7	1.3	206	3	AAW58830	Breast an
151	7	1.3	162	2	AAW81345	Aaw81345	[Delta-Hi	224	7	1.3	206	3	AAW58830	Breast an
152	7	1.3	162	2	AAW81347	Aaw81347	[Delta-Gl	225	7	1.3	207	2	AAW05800	Tilapia (
153	7	1.3	163	2	AAW80825	Aaw80825	Human MGD	226	7	1.3	207	2	AAW08122	Tilapia g
154	7	1.3	163	2	AAW81342	Aaw81342	Thrombopo	227	7	1.3	207	5	ABW89411	Human pol
155	7	1.3	163	2	AAW81354	Aaw81354	[Alal, Va	228	7	1.3	208	2	AAR97767	Human thr
156	7	1.3	163	2	AAW81373	Aaw81373	[Arg59]TP	229	7	1.3	213	2	AAR97773	Human thr
157	7	1.3	163	2	AAW81360	Aaw81360	[Alal, Va	230	7	1.3	216	2	AAW88434	Human MGD
158	7	1.3	163	2	AAW81359	Aaw81359	[Alal, Va	231	7	1.3	217	2	ADC31328	Human nov
159	7	1.3	163	2	AAW81366	Aaw81366	[Alal, Va	232	7	1.3	217	8	ADW09364	Human pro
160	7	1.3	163	2	AAW81371	Aaw81371	[Leu146]T	233	7	1.3	218	7	ADW88883	Enterococ
161	7	1.3	163	2	AAW81374	Aaw81374	[Arg115]T	234	7	1.3	222	6	ABW68944	Photorhab
162	7	1.3	163	2	AAW81369	Aaw81369	[Arg143]T	235	7	1.3	223	3	AAW56768	Human pro
163	7	1.3	163	2	AAW81363	Aaw81363	[Alal, Va	236	7	1.3	223	5	AAW15070	Humicola
164	7	1.3	163	2	AAW81368	Aaw81368	[Arg133]T	237	7	1.3	223	5	AAW80602	Humicola
165	7	1.3	163	2	AAW81362	Aaw81362	[Alal, Va	238	7	1.3	227	3	ADC32999	Human nov
166	7	1.3	163	2	AAW81367	Aaw81367	[Arg129]T	239	7	1.3	227	3	AAW09434	Arabidops
167	7	1.3	163	2	AAW81370	Aaw81370	[Leu82]TP	240	7	1.3	227	4	AAW08052	Arabidops
168	7	1.3	163	2	AAW81365	Aaw81365	[Alal, Va	241	7	1.3	227	4	ABG16682	Novel hum
169	7	1.3	163	2	AAW81361	Aaw81361	[Alal, Va	242	7	1.3	228	5	AAW15062	Endogluc
170	7	1.3	163	2	AAW97769	Aaw97769	Human thr	243	7	1.3	228	8	ADL05112	M. catarr
171	7	1.3	163	5	ABB07002	Abb07002	Human c-m	244	7	1.3	232	2	AAR81343	Thrombopo

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2005, 03:19:49 ; Search time 4589.43 Seconds
(without alignments)
4105.478 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 2617

Sequence: 1 MKSPVSLAPSGDEGSDRT.....LPAKGEVQNAISDQHGRN 495

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09776865/runat_08072005_175549_23379/app_query.fasta_1.1358
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09776865@cgn_1_1_6461/runat_08072005_175549_23379 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQ=0 -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPEXT=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsa1:*

9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2329	89.0	3189	3	CR618872	full-leng
2	2236	85.4	3202	3	AK029102	Mus muscu
3	1434	54.8	1051	1	AL550137	AL550137
4	1320.5	50.5	1059	5	BX425026	BX425026
5	1310	50.1	1038	5	BX439809	BX439809
6	1291	49.3	737	7	CR833737	CR833737
7	1267	48.4	721	7	CR833565	CR833565
8	1236	47.2	947	7	CO579484	ILLUMIGEN
9	1128	43.1	773	1	AJ734267	AJ734267

10	1118	42.7	740	6	CB317739	CB317739
11	1108	42.3	730	6	CA3227442	AGENCOURT
12	1075.5	41.1	901	6	CD106410	UI-M-FY0-
13	1055	40.3	708	7	CR763802	AGENCOURT
14	1038	39.7	801	4	BI661062	DKF2P4690
15	1001	38.2	813	7	CF618610	AGENCOURT
16	1000	38.2	605	5	BX479639	DKF2P686F
17	996	38.1	770	7	CF289959	AGENCOURT
18	988.5	37.8	791	7	CV110989	AGENCOURT
19	986	37.7	913	5	BX348297	AGENCOURT
20	971.5	37.1	831	5	AK087395	AGENCOURT
21	964	36.8	4078	3	AK045409	Mus muscu
22	953	36.4	754	4	BI907284	AGENCOURT
23	949	36.3	581	5	BP297030	AGENCOURT
24	948	36.2	570	5	BP297030	AGENCOURT
25	945	36.1	663	7	CV023522	AGENCOURT
26	942.5	36.0	776	5	BX882862	AGENCOURT
27	928.5	35.5	2014	3	AK082743	AGENCOURT
28	921	35.2	626	2	BP676817	AGENCOURT
29	915	35.0	711	4	BI697765	AGENCOURT
30	903	34.5	582	5	BF302998	AGENCOURT
31	903	34.5	630	6	CB138761	AGENCOURT
32	903	34.4	1531	3	BC023331	AGENCOURT
33	900	34.4	649	1	AJ734277	AGENCOURT
34	890	34.0	611	1	AJ734276	AGENCOURT
35	885	33.8	611	1	AJ734276	AGENCOURT
36	884.5	33.8	4630	3	AK040717	AGENCOURT
37	879.5	33.6	736	4	BI860521	AGENCOURT
38	877	33.5	700	7	CN786597	AGENCOURT
39	874.5	33.4	1494	9	AY407679	AGENCOURT
40	873	33.4	616	5	BP238694	AGENCOURT
41	868	33.2	895	2	BF539146	AGENCOURT
42	866	33.1	537	6	CB158910	AGENCOURT
43	866	33.1	537	6	CB158929	AGENCOURT
44	864.5	33.0	1485	9	AY407681	AGENCOURT
45	859	32.8	618	7	CF170114	AGENCOURT

ALIGNMENTS

RESULT 1
LOCUS CR618872
DEFINITION full-length cDNA clone CS0DI040YK17 of Placenta Cot 25-normalized
3189 bp mRNA linear HTC 21-JUL-2004
of Homo sapiens (human)

ACCESSION CR618872
VERSION CR618872.1 GI:50499679

KEYWORDS HTC; CDS; EST; cDNA

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600

REFERENCE Genoscope.
AUTHORS Direct Submission

TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :

JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES Location/Qualifiers

source 1..3189

/organism="Homo sapiens"

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSOD1040YK17"
 /tissue_type="Placenta Cot 25-normalized"
 /plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.: 3,97e-237 Length: 3189
 Score: 2329.00 Matches: 427
 Percent Similarity: 94.14% Conservative: 39
 Best Local Similarity: 86.26% Mismatches: 29
 Query Match: 89.00% Indels: 0
 DB: 3 Gaps: 0

US-09-776-865-4 (1-495) x CR618872 (1-3189)

Qy 1 MetLysSerProValSerAspLeuAlaProSerAspGlyGluGluGlySerAspArgThr 20
 Db 76 ATGAGGTCTCCGGTTCGAGACCTGCCCGGAACGATGCGAGGAGCAGCAGCGCAGC 135
 Qy 21 ProLeuLeuGlnArgAlaProArgAlaGluProAlaProValCysCysSerAlaArgThr 40
 Db 136 CCTCTTCTACCGGGCGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 195
 Qy 41 AsnLeuAlaPheLeuSerPhePheGlyPheValLeuLeuTyrSerLeuArgValAsnLeu 60
 Db 196 AACTTAGCAATTTGGCCCTTTTGGTTCCTTCATTGGTATGATGATGATGATGATGAT 255
 Qy 61 SerValAlaLeuValAspMetValAspSerLeuThrAlaLysAspAsnArgThrSer 80
 Db 256 AGTGTGGGTAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 315
 Qy 81 TyrGluCysAlaGluHisSerAlaProLeuValLeuHisAsnGlnThrGlyLysLys 100
 Db 316 AAGCGGTGTCAGACATTCGTCTCCCATAAAGTTCATCAATCAACCGGGAAGAAG 375
 Qy 101 TyrArgTyrAspAlaGluThrGlnGlyTyrPheLeuGlySerPhePheTyrGlyTyrile 120
 Db 376 TACCAATGGGTGACAGAACTCAAGATGATGATGATGATGATGATGATGATGATGAT 435
 Qy 121 IleThrGlnIleProGlyGlyTyrValAlaSerArgSerGlyGlyLysLeuLeuGly 140
 Db 436 ATCACACAGATTCCTGGAGGATATGTCGACAAATAGGGGGAATAGGCTGTAGGA 495
 Qy 141 PheGlyIlePheAlaThrAlaIlePheThrLeuPheThrProLeuAlaAlaaspPheGly 160
 Db 496 TTTGGGATCCTTGGCAGCTGCTGCTCCCTGTTCTCCTCCATGCTGCGAGATTTAGGA 555
 Qy 161 ValGlyAlaLeuValAlaLeuArgAlaLeuGluGlyLeuGlyGluGlyValThrTyrPro 180
 Db 556 GTTGGACCACTCATGTACTCAGACACTAGAGGAGCTAGGAGGAGGTTTACATTTCCA 615
 Qy 181 AlaMetHisAlaMetTyrSerSerTrpAlaProProLeuGluArgSerLysLeuLeuSer 200
 Db 616 GCCATGATGCATGCTGCTTCTTGGGCTCCCTCTTGAAGAAGCAAACTCTTAGC 675
 Qy 201 IleSerTyrAlaGlyAlaGluLeuGlyThrValValSerLeuProLeuSerGlyValile 220
 Db 676 ATTTATATGACGAGGACAGCTGGGACAGTAATTTCTTCTCTCTTCTTCTGGATAAAT 735
 Qy 221 CysTyrTyrMetAsnTrpThrTyrValPheTyrPhePheGlyIleValGlyIleIleTrp 240
 Db 736 TGTACTATATGAATGGACTATGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 795
 Qy 241 PheIleLeuTrpIleCysLeuValSerAspThrProGluThrHisLysThrIleThrPro 260
 Db 796 TTTCTTTTGTGATCTGTTAGTACACACACACACACACACACACACACACACACACAC 855
 Qy 261 TyrGluLysGluTyrIleLeuSerSerLeuLysAsnGlnLeuSerSerGlnLysSerVal 280
 Db 856 TATGAAAGGAAGGATACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 915

Qy 281 ProTrpIleProMetLeuLysSerLeuProLeuTrpAlaIleValValAlaHisPheSer 300
 Db 916 CCGTGGGTACCAATTTTAAAAATCCCTGCCACTTTGGGTATCGTAGTTGCACACTTTCT 975
 Qy 301 TyrAsnTrpThrPheTyrThrLeuLeuThrLeuLeuProThrTyrMetLysGluValLeu 320
 Db 976 TACAACGGGACTTTTATATCTTTTATGACATTTATGCTTCTTATATGAAGAGATCTTA 1035
 Qy 321 ArgPheAsnIleGlnGluAsnGlyPheLeuSerAlaValProTyrLeuGlyCysTrpLeu 340
 Db 1036 AGTTTCATGTTCAAGAGATGGGTTTATCTTCATTCCTTATTTAGGCTCTTGGTTA 1095
 Qy 341 CysMetIleLeuSerGlyGlnAlaAlaAspAsnLeuArgAlaArgTrpAsnPheSerThr 360
 Db 1096 TGTATGATCCTGCTCGTCAAGCTGCTGACAAATTAAGGGCAAAATGGAATTTTCACT 1155
 Qy 361 LeuTrpValArgArgValPheSerLeuIleGlyMetIleGlyProAlaIlePheLeuVal 380
 Db 1156 TTATGTGTTCGAGAAATTTTATAGCCTTATAGGAATGATTTGGACCTGAGTATTCCTG 1215
 Qy 381 AlaAlaGlyPheIleGlyCysAspTyrSerLeuAlaValAlaPheLeuThrIleSerThr 400
 Db 1216 GCTCTGGCTTCATGCTGCTGATATCTTTGGCCGTGCTTCTTAACATATATCAACA 1275
 Qy 401 ThrLeuGlyGlyPheCysSerSerGlyPheSerIleAsnHisLeuAspIleAlaProSer 420
 Db 1276 ACCTGGGAGGCTTTGCTCTCTGATTTAGCATCAACCATCTGGATATTTGCTCTCTTC 1335
 Qy 421 TyrAlaGlyIleLeuLeuGlyIleThrAsnThrPheAlaThrIleProGlyMetIleGly 440
 Db 1336 TATCTGGTATCTCTCGGGCATCACAAATACATTTGCCACTATTCAGGAATGTTGGG 1395
 Qy 441 ProIleAlaAlaArgSerLeuThrProGluAsnThrIleGlyGluTrpGlnThrValPhe 460
 Db 1396 CCCGTCATGCTAAAGTCTGACCCCTGATPACACTGTGGAGATGGCAACCGTGTTC 1455
 Qy 461 CysIleAlaAlaIleAsnValPheGlyAlaIlePheThrLeuPheAlaLysGly 480
 Db 1456 TATATGCTGCTGCTATTAATGTTTTTGGTGCCATTTCTTTACACTATTTGCCAAAGGT 1515
 Qy 481 GluValGlnAsnTrpAlaIleSerAspHisGlnGlyHisArgAsn 495
 Db 1516 GAAGTACAAAACCTGGCTCTCAATGATCACCATGCACACAGACAC 1560

RESULT 2
 LOCUS AK029102 3202 bp mRNA linear HTC 03-APR-2004
 DEFINITION Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:473491M05 product:SIALIN homolog [Homo sapiens], full insert sequence.
 ACCESSION AK029102
 VERSION AK029102.1 GI:26325087
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2005, 07:16:14 ; Search time 840.684 Seconds
(without alignments)
3696.712 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 2617

Sequence: 1 MKSPVSLAPSDGEGSDRT.....LPKAGEVQNWAIHQHGRN 495

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6330943 seqs, 3139157217 residues

Total number of hits satisfying chosen parameters: 12661886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_epool/US09776865/runat 08072005_175551_23472/app_query.fasta_1.1358
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLCN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -UNAP=US09776865@cgn_1_1065@runat 08072005_175551_23472
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/FCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/FCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2617	100.0	2844	9	US-09-776-865-3	Sequence 3, Appli
2	2617	100.0	2844	21	US-10-823-506-3	Sequence 3, Appli
3	2329	89.0	2602	21	US-10-823-506-1	Sequence 1, Appli
4	2329	89.0	2626	14	US-10-198-846-12482	Sequence 12482, A
5	2329	89.0	2930	9	US-09-776-865-1	Sequence 1, Appli
6	2329	89.0	2930	21	US-10-823-506-7	Sequence 7, Appli
7	2329	89.0	3292	21	US-10-887-553A-858	Sequence 858, App
8	2329	89.0	3329	19	US-10-755-889-587	Sequence 587, App
9	2329	89.0	3362	10	US-09-814-353-19097	Sequence 19097, A
10	2322	88.7	1488	13	US-10-098-841-322	Sequence 322, App
11	2221	84.9	1485	21	US-10-823-506-9	Sequence 9, Appli
12	1971	75.3	1485	21	US-10-823-506-11	Sequence 11, Appli
13	1386	53.0	1872	17	US-10-264-049-302	Sequence 302, App
14	1380	52.7	1651	17	US-10-264-237-946	Sequence 946, App
15	980	37.4	1811	9	US-09-740-041-1	Sequence 1, Appli
16	980	37.4	1811	16	US-10-389-967-1	Sequence 1, Appli
17	979	37.4	1767	22	US-10-499-731-29	Sequence 29, Appli
18	979	37.4	1850	22	US-10-499-731-45	Sequence 45, Appli
19	979	37.4	3838	21	US-10-887-553A-1200	Sequence 1200, Ap
20	969.5	37.0	1803	22	US-10-499-731-47	Sequence 47, Appli
21	969.5	37.0	1822	22	US-10-499-731-58	Sequence 58, Appli
22	969	37.0	2528	20	US-10-734-731-13	Sequence 13, Appli
23	969	37.0	2528	20	US-10-807-500-13	Sequence 13, Appli
24	968	37.0	3982	18	US-10-205-331-8	Sequence 8, Appli
25	968	37.0	3982	20	US-10-734-731-11	Sequence 11, Appli
26	968	37.0	3982	20	US-10-807-500-11	Sequence 11, Appli
27	963.5	36.8	3946	20	US-10-734-731-9	Sequence 9, Appli
28	963.5	36.8	3946	21	US-10-807-500-9	Sequence 9, Appli
29	963.5	36.8	3946	21	US-10-887-553A-1016	Sequence 1016, Ap
30	958	36.6	1701	18	US-10-287-226-299	Sequence 299, App
31	957.5	36.6	1701	16	US-10-233-045-21	Sequence 21, Appli
32	956.5	36.5	2607	9	US-09-915-181A-1	Sequence 1, Appli
33	933.5	35.7	2024	18	US-10-152-319A-2039	Sequence 2039, Ap
34	933.5	35.7	2024	20	US-10-734-731-5	Sequence 5, Appli
35	933.5	35.7	2024	20	US-10-807-500-5	Sequence 5, Appli
36	932	35.6	2366	20	US-10-734-731-1	Sequence 1, Appli
37	932	35.6	2366	20	US-10-807-500-1	Sequence 1, Appli
38	931	35.6	2716	17	US-10-314-790-6	GENERAL INFORMA
39	931	35.6	2716	20	US-10-734-731-3	Sequence 3, Appli
40	931	35.6	2716	20	US-10-807-500-3	Sequence 3, Appli
41	927	35.4	2836	20	US-10-734-731-7	Sequence 7, Appli
42	927	35.4	2836	20	US-10-807-500-7	Sequence 7, Appli
43	895	34.2	1975	18	US-10-296-115-461	Sequence 461, App
44	800.5	30.6	1700	17	US-10-388-934-516	Sequence 516, App
45	800.5	30.6	1700	18	US-10-152-319A-2122	Sequence 2122, Ap

ALIGNMENTS

RESULT 1

US-09-776-865-3
; Sequence 3, Application US/09776865
; Patent No. US20020061846A1
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; TITLE OF INVENTION: Methods for Preventing or Attenuating Pathoangiogenic Conditions
; FILE REFERENCE: 22100-0100 46126-252687
; CURRENT APPLICATION NUMBER: US/09/776,865
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 60/179,870
; PRIOR FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 2844


```
; TYPE: DNA
; ORGANISM: Ovis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)..(1568)
US-09-776-865-3

Alignment Scores:
Pred. No.: 6,74e-271 Length: 2844
Score: 2617.00 Matches: 495
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-776-865-4 (1-495) x US-09-776-865-3 (1-2844)

Qy 1 MetLysSerProValSerAspLeuAlaProSerAspGlyGluGluGlySerAspArgThr 20
Db 84 ATGAAGTCCCGGTTTCGGACTTAGCCCGAGCGACGGCGAGGAGGCTCGGACCGCACA 143
Qy 21 ProLeuLeuGlnArgAlaProArgAlaGluProAlaProValCysCysSerAlaArgTyr 40
Db 144 CGCTCTCGACGGCGCCCGCGGGCGGAACCGCTCCAGTATGCTCTCTCGCTGTAC 203
Qy 41 AsnLeuAlaPheLeuSerPhePheGlyPhePheValLeuTyrSerLeuArgValAsnLeu 60
Db 204 AACCTAGCATTTTGTCTCTTTTGGTTCTTCGTTCTCTATTATTCATTCAGGCTGAATCG 263
Qy 61 SerValAlaLeuValAspMetValAspSerAsnThrThrAlaLysAspAsnArgThrSer 80
Db 264 AGCGTTGCACCTAGTGGACATGCTGGATTCAAAACACACTGCGCAAGATAATAGAACGTC 323
Qy 81 TyrGluCysAlaGluHisSerAlaProIleLysValLeuHisAsnGlnThrGlyLysLys 100
Db 324 TACGAGTGTGACAGACATCTGCTCCCATAAAGTCTTCACAAACAAACGGGTAAGAAAG 383
Qy 101 TyrArgTyrAspAlaGluThrGlnGlyThrIleLeuGlySerPhePheTyrGlyTyrIle 120
Db 384 TACCGGTGGGATGCAGAACTCAAGATGGATTCTCGGATCTTTTCTATGGCTACATC 443
Qy 121 IleThrGlnIleProGlyGlyTyrValAlaSerArgSerGlyGlyLysLeuLeuGly 140
Db 444 ATCACAAATTCCTGGAGGATATGTTGCCAGAGAGTGGGGGGAAGCTGTGCTAGGA 503
Qy 141 PheGlyIlePheAlaThrAlaIlePheThrLeuPheThrProLeuAlaAlaAspPheGly 160
Db 504 TTCGGGATCTTTGTACAGCTATCTTACCCTGTTCACCTCCCTCGCTGCGAGATTTCGA 563
Qy 161 ValGlyAlaLeuValAlaLeuArgAlaLeuGluGlyLeuGlyGluGlyValThrTyrPro 180
Db 564 GTCGAGGCCCTTTGTCACCTCAGGCGCACTAGAGGCTTAGGAGGGGTGTACATATCCA 623
Qy 181 AlaMetHisAlaMetTyrSerSerTyrAlaProProLeuGluArgSerLysLeuLeuSer 200
Db 624 GCCATGATGCATGTGTCTTCATGGGCTCCCTCTTCTTGAAGAAGCAAGCTTCGAGT 683
Qy 201 IleSerTyrAlaGlyAlaGlnLeuGlyThrValValSerLeuProLeuSerGlyValIle 220
Db 684 ATTTATATGACGAGGACAACTTGGGACAGTAGTTCTCTCTCTCTTCTGAGGATTAAT 743
Qy 221 CysTyrTyrMetAsnTyrThrTyrValPheTyrPhePheGlyIleValGlyIleIleTyr 240
Db 744 TGCTACTATATGAATTGGACTATGCTCTTCTATTCTTTTGGCAITGTTGGAATCATCTGG 803
Qy 241 PheIleLeuTyrIleCysLeuValSerAspThrProGluThrHisLysThrIleThrPro 260
Db 804 TTTATTTTATGGATCTGCTTAGTGTATGATACACAGAACTCACAAGACAATCACTCCC 863
Qy 261 TyrGluLysGluTyrIleLeuSerSerLeuLysAsnGlnLeuSerSerGlnLysSerVal 280
Db 864 TATGAAAGGAGTATATCTTTTCATCATTAATAAATACAGCTCTCTTTCAGAAAGTCAGTG 923
```

```
Qy 281 ProTyrIleProMetLeuLysSerLeuProLeuTyrProAlaIleValValAlaHisPheSer 300
Db 924 CCGTGGATACCTATGCTGAAATCACTGCCACTTTGGGCTATTGTCGTTGCACATTTTCT 983
Qy 301 TyrAsnTyrPheTyrThrLeuLeuThrLeuLeuProThrTyrMetLysGluValLeu 320
Db 984 TACAACCTGGACTTTTATATCTTTTGTGACCTTATTTGCTACTTACATGAAGGAAGTCTTA 1043
Qy 321 ArgPheAsnIleGlnGluAsnGlyPheLeuSerAlaValProTyrLeuGlyCysTyrLeu 340
Db 1044 AGTTTCATATTCAAGAGATGGGTTTTTATTCGAGTCCCTTATTTAGGTTGTGGTTA 1103
Qy 341 CysMetIleLeuSerGlyGlnAlaAlaAspAsnLeuArgAlaArgTyrAsnPheSerThr 360
Db 1104 TGTATGATCTGTCGGGTCAAGCTGCTGCAATTTAAGGGCAAGATGGAATTTTCAACT 1163
Qy 361 LeuTyrValArgValPheSerLeuIleGlyMetIleGlyProAlaIlePheLeuVal 380
Db 1164 CTGTGGGTTTCAAGAGATTTTATAGCCTTATAGGGAATGATTCGACCTGCGATATTCTCGTT 1223
Qy 381 AlaAlaGlyPheIleGlyCysAspTyrSerLeuAlaValAlaPheLeuThrIleSerThr 400
Db 1224 GCCGACGATTTATAGGCTGTGATTATCTTGGCTGTTCATCTTCAACATATCAACA 1283
Qy 401 ThrLeuGlyGlyPheCysSerSerGlyPheSerIleAsnHisLeuAspIleAlaProSer 420
Db 1284 ACCCTGGAGGCTTTTGTCTCTCTGGATTTAGCATCAACCATCTGGACATTTGCTCTCTCG 1343
Qy 421 TyrAlaGlyIleLeuLeuGlyIleThrAsnThrPheAlaThrIleProGlyMetIleGly 440
Db 1344 TATGCTGATTTCTCTCGGCATCACAAATACCTTTGCCACTATTTCTGGAATGATTGGG 1403
Qy 441 ProfileAlaArgSerLeuThrProGluAsnThrIleGlyGluTyrGlnThrValPhe 460
Db 1404 CCATCATTCGACAGAGTCTTACCCCTGAGACACTATTGAGAGATGCGCAACTGTTTC 1463
Qy 461 CysIleAlaAlaIleAsnValPheGlyValaIlePhePheThrLeuPheAlaLysGly 480
Db 1464 TGCATCGCTGCTGCTATCAATGATTTTGTGTCATTTTCTTTCACACTATTTCGCCAAGGT 1523
Qy 481 GluValGlnAsnTyrAlaIleSerAspHisGlnGlyHisArgAsn 495
Db 1524 GAAGTCGAAACTGGGCCCATCAGTGATCACCAAGGACACAGAAAC 1568

RESULT 2
US-10-823-506-3
; Sequence 3, Application US/10823506
; Publication No. US20050002931A1
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/01US
; CURRENT APPLICATION NUMBER: US/10/823,506
; PRIOR FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: US/09/359,167
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60-693,843
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.0
; LENGTH: 2844
; TYPE: DNA
; ORGANISM: Ovis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)..(1568)
US-10-823-506-3

Alignment Scores: 6,74e-271 Length: 2844
Pred. No.: 2617.00 Matches: 495
Score:
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2005, 06:18:14 ; Search time 233.817 Seconds

(without alignments)

3464.068 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 2617

Sequence: 1 MKSPVSLAPSGDGERGSDT.....LFAKGEVQNWAIHQHGRN 495

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09776865/runat_08072005_175549_23397/app_query.fasta_1.1358
-DB=Issued Patents NA -QFWT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09776865@cgn_1_1_277@runat_08072005_175549_23397 -NCPUS=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2617	100.0	2844	US-09-359-167-3	Sequence 3, Appli
2	2329	89.0	2513	US-09-949-016-1834	Sequence 1834, Ap
3	2329	89.0	2602	US-09-359-167-1	Sequence 1, Appli
4	2329	89.0	2930	US-09-359-167-7	Sequence 7, Appli
5	2221	84.9	1485	US-09-359-167-9	Sequence 9, Appli
6	1971	75.3	1485	US-09-359-167-11	Sequence 11, Appli
7	980	37.4	1811	US-09-740-041-1	Sequence 1, Appli
8	956.5	36.5	2607	US-09-915-181A-1	Sequence 1, Appli
9	932	35.6	2366	US-09-949-016-5483	Sequence 5483, Ap
10	931	35.6	2716	US-08-647-484-1	Sequence 1, Appli
11	931	35.6	2716	US-08-647-484-3	Sequence 3, Appli
12	931	35.6	2716	US-08-647-481-1	Sequence 1, Appli

13	931	35.6	2716	1	US-08-647-481-3	Sequence 3, Appli
14	931	35.6	2716	1	US-08-430-033A-1	Sequence 1, Appli
15	931	35.6	2716	1	US-08-430-033A-3	Sequence 3, Appli
16	931	35.6	2716	5	PCT-US96-05792-1	Sequence 1, Appli
17	931	35.6	2716	5	PCT-US96-05792-3	Sequence 3, Appli
18	874.5	33.4	2626	4	US-09-949-016-745	Sequence 745, App
19	761	29.1	2612	4	US-09-270-767-13970	Sequence 13970, A
20	753	28.8	2266	2	US-08-724-394A-18	Sequence 18, Appl
21	753	28.8	2270	4	US-09-949-016-5577	Sequence 5577, Ap
22	738	28.2	1229	4	US-09-023-655-660	Sequence 660, App
23	621.5	23.7	2296	4	US-09-949-016-1182	Sequence 1182, Ap
24	605.5	23.1	1643	2	US-08-805-118-2	Sequence 2, Appli
25	605.5	23.1	1643	3	US-09-391-958-2	Sequence 2, Appli
26	605.5	23.1	1780	2	US-08-724-394A-19	Sequence 19, Appl
27	450	17.2	1383	4	US-09-489-039A-2492	Sequence 2492, Ap
28	442	16.9	1699	4	US-09-799-451-815	Sequence 815, App
29	428.5	16.4	1074	4	US-09-270-767-13957	Sequence 13957, A
30	398	15.2	3486	4	US-09-976-594-1018	Sequence 1018, Ap
31	389	14.9	976	4	US-09-270-767-13814	Sequence 13814, A
32	385.5	14.7	1413	4	US-09-328-352-2786	Sequence 2786, Ap
33	385.5	14.0	1368	4	US-09-489-039A-2771	Sequence 2771, Ap
34	365	13.9	1317	3	US-09-172-952-5	Sequence 5, Appli
35	365	13.9	1500	4	US-09-922-501-11	Sequence 11, Appl
36	365	13.9	9334	3	US-09-172-952-19	Sequence 19, Appl
37	363	13.9	63783	4	US-09-949-016-13576	Sequence 13576, A
38	360	13.8	471	4	US-09-270-767-30032	Sequence 30032, A
39	357	13.6	217	4	US-09-016-434-602	Sequence 602, App
40	353	13.5	1305	4	US-09-489-039A-6462	Sequence 6462, Ap
41	341	13.0	492	4	US-09-270-767-1297	Sequence 1297, Ap
42	341	13.0	492	4	US-09-270-767-16579	Sequence 16579, A
43	338.5	12.9	1269	4	US-09-891-641-65	Sequence 65, Appl
44	326.5	12.5	1380	4	US-09-489-039A-3459	Sequence 3459, Ap
45	318	12.2	1600	4	US-09-922-501-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-359-167-3
; Sequence 3, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/0105
; CURRENT APPLICATION NUMBER: US/09/359,167
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: 60-693,843
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2844
; TYPE: DNA
; ORGANISM: Ovis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)..(1568)
US-09-359-167-3

Alignment Scores:
Pred. No.: 9,11e-302 Length: 2844
Score: 2617.00 Matches: 495
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-776-865-4 (1-495) x US-09-359-167-3 (1-2844)

Qy 1 MetLysSerProValSerAspLeuAlaProSerAspGlyGluGlySerAspArgThr 20

Db 84 ATGAAGTCCCGGTTTCGGACCTTAGCCCGGAGCGAGCGGAGGGCTCGGACCGCAC 143
Qy 21 ProLeuLeuGlnArgAlaProArgAlaGluProAlaProValCysCysSerAlaArgTyr 40
Db 144 CGCTCTCGACGGCGCGCGCGGCGGACCGCTCCAGTATGCTGCTCTGCTGTAC 203
Qy 41 AsnLeuAlaPheLeuSerPheGlyPhePheValLeuTyrSerLeuArgValAsnLeu 60
Db 204 AACCTAGCATTTCCTCTCTTTTGGTTCCTCTCTCTCTCTCTCTCTCTCTCTCT 263
Qy 61 SerValAlaLeuValAspMetValAspSerAsnThrThrAlaLysAspAsnArgThrSer 80
Db 264 AGCGTGTGACATAGTGAGATGCTGATTCACCACTGCCAAGATATAAGACGTC 323
Qy 81 TyrGluCysAlaGluHisSerAlaProIleLysValLeuHisAsnGlnThrGlyLysLys 100
Db 324 TAGGAGTGTGACAGCATTCCTGCTCCCATAAAGTCTTCACACCAACCGGTAAAG 383
Qy 101 TyrArgTTPAspAlaGluThrGlnGlyTrpIleLeuGlySerPhePheTyrGlyTyrIle 120
Db 384 TACCGGTGGGATGCAGAACTCAAGATGGATTCCTCGGATCTTTTCTATGGCTACATC 443
Qy 121 IleThrGlnIleProGlyGlyTyrValAlaSerArgSerGlyGlyLysLeuLeuGly 140
Db 444 ATCACACAAATTCCTGGAGGATATGTTGCCAGCAGAGTGGGGGAAGCTGTGCTAGGA 503
Qy 141 PheGlyIlePheAlaThrAlaIlePheThrLeuPheThrProLeuAlaAlaAspPheGly 160
Db 504 TTCGGGATCTTGTACAGCTATCTTCACCTGTTCACTCCCTCGCTGCAGATTTCGA 563
Qy 161 ValGlyAlaLeuValAlaLeuArgAlaLeuGluGlyLeuGlyGluGlyValThrTyrPro 180
Db 564 GTCCGAGCCCTGTTGCACTCAGGCACTAGAAAGGCTAGGAGAGGGTGCACATATCCA 623
Qy 181 AlaMetHisAlaMetTyrSerSerTyrAlaProProLeuGluUArgSerLysLeuLeuSer 200
Db 624 GCCATGCAATGCATGTGTCTTCATGGGCTCCCCCTCTGAAAGAACCAAGCTTCAGT 683
Qy 201 IleSerTyrAlaGlyAlaGlnLeuGlyThrValValSerLeuProLeuSerGlyValIle 220
Db 684 ATTTCAATATGCGAGGACACACTGGGACAGTAGTTCTCTCTCTCTCTCTCTCTCT 743
Qy 221 CysTyrTyrMetAsnTyrThrTyrValPheTyrPhePheGlyIleValGlyIleIleTyr 240
Db 744 TGCTACTATATGAATTGGACTATGCTCTCTATTTCTTGGCATGTTTGAATCATCTGG 803
Qy 241 PheIleLeuTyrIleCysLeuValSerAspThrProGluThrHisLysThrIleThrPro 260
Db 804 TTTATTTTATGGATCTGCTTAGTTAGTATACACCAAGAACTCACAAAGCAATCACTCCC 863
Qy 261 TyrGluLysGluTyrIleLeuSerSerLeuLysAsnGlnLeuSerSerGlnLysSerVal 280
Db 864 TATGAAAGGAGTATATTTCTTTCATCATATAAATAATCAGCTCTCTTCACAGAGTCAGT 923
Qy 281 ProTyrIleProMetLeuLysSerLeuProLeuTyrAlaIleValValAlaHisPheSer 300
Db 924 CGTGGATACCTATGCTGAAATCACTGCCACTTTGGGCTATTGCTGTTGCACATTTTCT 983
Qy 301 TyrAsnTyrThrPheTyrThrLeuLeuThrLeuLeuProThrTyrMetLysGluValLeu 320
Db 984 TACAACTGGACTTTTATACCTTTGTGACCTTATTTGCCCTACTTACATGAAGGAATCCTA 1043
Qy 321 ArgPheAsnIleGlnGluAsnGlyPheLeuSerAlaValProTyrLeuGlyCysTyrLeu 340
Db 1044 AGTTCAATATTCAGAGATGGGTTTTTATCTGAGTCCCTTATTTAGTTTGTGTTA 1103
Qy 341 CysMetIleLeuSerGlyGlnAlaAlaAspAsnLeuArgAlaArgTyrAsnPheSerThr 360
Db 1104 TGTATGATCTGTCCGCTCAAGCTGCTGACAAATTAAGGGCAAGATGAAATTTTCAACT 1163
Qy 361 LeuTyrValArgArgValPheSerLeuIleGlyMetIleGlyProAlaIlePheLeuVal 380
Db 1164 CTGTGGGTTCGAAGAGATTTTAGCCTTATAGGATGATTTGGACCTGCGATATTCCTGGT 1223

Qy 381 AlaAlaGlyPheIleGlyCysAspTyrSerLeuAlaValAlaPheLeuThrIleSerThr 400
Db 1224 GCCGAGGATTTATAGCTGTGATTAATTCCTTGGCTGTTCATCTTAACCATATCAACA 1283
Qy 401 ThrLeuGlyGlyPheCysSerSerGlyPheSerIleAsnHisLeuAspIleAlaProSer 420
Db 1284 ACCCTGGAGGCTTTTGTCTCTCTCTGATTTAGCATCAACCATCTGGACATTCCTCT 1343
Qy 421 TyrAlaGlyIleLeuLeuGlyIleThrAsnThrPheAlaThrIleProGlyMetIleGly 440
Db 1344 TATCTGATATTCCTCTGGGATCACAAATACCTTTGCCACTATTCCTGGAAATGATGGG 1403
Qy 441 ProIleAlaAlaArgSerLeuThrProGluAsnThrIleGlyGluTyrGlnThrValPhe 460
Db 1404 CCATCATTCGCAGAGTCTTACCCCTGAGACACATATGGAGAAATGGCAACTGTTTTC 1463
Qy 461 CysIleAlaAlaAlaIleAsnValPheGlyAlaIlePhePheThrLeuPheAlaLysGly 480
Db 1464 TGCATCGCTGCTGCTATCAATGATTTTGGTGCCATTTTCTTCACTATTCGCCAAAGT 1523
Qy 481 GluValGlnAsnTyrAlaIleSerAspHisGlnGlyHisArgAsn 495
Db 1524 GAAGTGCAAAACTGGGCTATCAGTGATCACCAAGACACAGAAAC 1568

RESULT 2

US-09-949-016-1834
; Sequence 1834, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1834
; LENGTH: 2513
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1834

Alignment Scores:

Pred No.:	1,96e-267	Length:	2513
Score:	2329.00	Matches:	427
Percent Similarity:	94.14%	Conservative:	39
Best Local Similarity:	86.26%	Mismatches:	29
Query Match:	89.00%	Indels:	0
DB:	4	Gaps:	0

US-09-776-865-4 (1-495) x US-09-949-016-1834 (1-2513)

Qy 1 MetLysSerProValSerAspLeuAlaProSerAspGlyGluGlySerAspArgThr 20
Db 272 ATGAGGCTCCGGTTCGAGACCTGGCCCGGAACGATGGCGAGGAGAGCAGCGACGACG 331
Qy 21 ProLeuLeuGlnArgAlaProArgAlaGluProAlaProValCysCysSerAlaArgTyr 40
Db 332 CCTCTTACGGCGCCGCCACCGGCGGAAGCCCTCCAGTGTCTCTCTGCTCGTTAC 391
Qy 41 AsnLeuAlaPheLeuSerPhePhePheValLeuTyrSerLeuArgValAsnLeu 60
Db 392 AACTTAGCAATTTGGCTTTTGGTTTCTTCACTGTGTATGATGATGATGATGATGATG 451
Qy 61 SerValAlaLeuValAspMetValAspSerAsnThrThrAlaLysAspAsnArgThrSer 80

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2005, 02:43:23 ; Search time 742.26 Seconds
(without alignments)
3947.771 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 2617

Sequence: 1 MKSPVSLAPDGESEGRD.....LFAKGEVQNWAISSHQGRN 495

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool-US09776865/runat_08072005_175548_23365/app_query.fasta_1.1358
-DB=N Geneseq_16Dec04 -QWTF=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09776865 @CGN_1_1004 @runat_08072005_175548_23365 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_16Dec04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2002bs.*
- 8: geneseqn2003as.*
- 9: geneseqn2003bs.*
- 10: geneseqn2003cs.*
- 11: geneseqn2003ds.*
- 12: geneseqn2004as.*
- 13: geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2617	100.0	2844	3 AAZ50876	AAZ50876 Sheep GBS
2	2617	100.0	2844	4 AAD10326	Aad10326 Sheep GBS
3	2329	89.0	2511	12 AD084158	Adg84158 Human tum
4	2329	89.0	2512	4 AAF55900	Aaf55900 Human AST
5	2329	89.0	2602	3 AAZ50875	AAZ50875 Partial h

6	2329	89.0	2626	11 ACN91332	ACN91332 Breast ca
7	2329	89.0	2930	3 AAZ50879	Aaz50879 Full leng
8	2329	89.0	2930	4 AAD10325	Aad10325 Human gro
9	2329	89.0	3329	12 AD075057	Adj75057 Marker ge
10	2329	89.0	3329	13 ADRI4586	Adri4586 Human NF-
11	2329	89.0	3329	13 ADP25216	Adp25216 PRO polyp
12	2329	89.0	3362	5 ADL45207	Adl45207 Human ova
13	2322	88.7	1488	4 AAI58115	Aai58115 Human pol
14	2302	88.0	2712	4 AAK94876	Aak94876 Human ful
15	2302	88.0	2712	12 ADL32035	Adl32035 Full leng
16	2221	84.9	1485	3 AAZ50880	Aaz50880 Human/She
17	1971	75.3	1485	3 AAZ50881	Aaz50881 Human/She
18	1578	60.3	2670	4 AAH79234	Aah79234 Human sod
19	1386	53.0	1872	6 ABQ54422	Abq54422 Human ova
20	1380	52.7	1651	6 ABL90384	AbL90384 Human pol
21	1217	46.5	929	4 AAI59901	Aai59901 Human pol
22	1104	42.2	853	4 AAK93901	Aak93901 Human CDN
23	1104	42.2	853	12 ADL30328	Adl30328 3' end of
24	1022.5	39.1	838	4 AAK92364	Aak92364 Human CDN
25	1022.5	39.1	838	12 ADL28791	Adl28791 5' end of
26	981.5	37.9	1939	4 ABL03769	AbL03769 Drosophil
27	985.5	37.7	1641	4 ABL18113	AbL18113 Drosophil
28	985.5	37.7	1786	4 ABL07417	AbL07417 Drosophil
29	981.5	37.5	4899	4 ABL03768	AbL03768 Drosophil
30	980	37.4	1811	6 ABK87049	Abk87049 Human tra
31	980	37.4	1811	10 ADG88328	Adg88328 Human tra
32	979	37.4	1767	10 ADD01393	Add01393 Human TCH
33	979	37.4	1850	10 ADD01409	Add01409 Human TCH
34	979	37.4	3671	9 AAL62505	Aal62505 Human tra
35	979	37.4	3952	13 ADR08065	Adr08065 Full leng
36	969.5	37.0	1803	10 ADD01411	Add01411 Mouse TCH
37	969.5	37.0	1822	10 ADD01422	Add01422 Mouse TCH
38	969	37.0	2528	8 ABX13555	Abx13555 Murine DN
39	969	37.0	2528	10 ADC15493	Adc15493 Mouse DNP
40	968	37.0	3982	8 ABX13554	Abx13554 Rat DNP1
41	968	37.0	3982	9 ACF25330	Acf25330 Rat Na-de
42	968	37.0	3982	10 ADC15491	Adc15491 Rat DNPI
43	963.5	36.8	3946	8 ABX13553	Abx13553 Human DNP
44	963.5	36.8	3946	10 ADC15489	Adc15489 Human DNP
45	958	36.6	1701	10 ADJ95071	Adj95071 Novel NOV

ALIGNMENTS

RESULT 1

AAZ50876
ID AAZ50876 standard; cDNA; 2844 BP.
XX
AC AAZ50876;
XX
DT 31-MAY-2000 (first entry)
XX
DE Sheep GBS toxin receptor (SP55) cDNA.

Sheep GBS toxin receptor; group B beta-haemolytic streptococci; SP55;
pathological vascularisation; cancer metastases; angiogenesis;
neovascularisation; reperfusion injury; scarring; keloid;
chronic inflammation; rheumatoid arthritis; psoriasis; neural injury;
endothelial cell proliferation; antibacterial; anticancer;
anti-angiogenic; anti-inflammatory; anti-arthritis; anti-psoriatic; ss.

Ovis sp.

Key Location/Qualifiers
CDS 84..1571
FT /*tag= a
FT /product= "Sheep GBS toxin receptor"

XX WO200005375-A1.

XX 03-FEB-2000.

XX 22-JUL-1999; 99WO-US016676.

XX PR 22-JUL-1998; 98US-0093843P.
XX PA (UYVA-) UNIV VANDERBILT.
XX PI Hellerqvist CG, Fu C;
XX WPI; 2000-205377/18.
DR P-PSDB; AAY45088.
XX New polynucleotide encoding mammalian receptor for streptococcus toxin,
PT useful for diagnosis and treatment of, e.g. pneumonia in neonates.
XX
PS Claim 3; Page 83-86; 109pp; English.
XX
CC The present cDNA sequence encodes partial sheep GBS (group B beta -
CC haemolytic streptococci) toxin receptor (SP55). This sequence was cloned
CC using a primary culture of sheep lung endothelial cells. Expression
CC vectors comprising this cDNA can be transformed into host cells to
CC express GBS toxin receptor and its fragments. Detecting the receptor in
CC tissues is used to diagnose pathological vascularisation, e.g. for
CC detecting cancer metastases. GBS toxin receptors are useful for treating
CC conditions associated with pathological angiogenesis or
CC neovascularisation (specifically cancer, reperfusion injury, scarring
CC during wound healing, keloids, chronic inflammation (rheumatoid arthritis
CC or psoriasis) or neural injury), and to raise specific antibodies used
CC for treating early onset disease. Inhibitors of this receptor are useful
CC for treating pathological or hypoxia-induced endothelial cell
CC proliferation and migration
XX
SQ Sequence 2844 BP; 745 A; 623 C; 594 G; 882 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,99e-256 Length: 2844
Score: 2617.00 Matches: 495
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-776-865-4 (1-495) x ANZ50876 (1-2844)

QY 1 MetLysSerProValSerAspLeuAlaProSerAspGlyGluGlySerAspArgThr 20
Db 84 ATGAAGTCCCGGTTTCGAGCTTAGCCCCGAGCGAGCGAGGCGGTCGGACCGCAC 143

QY 21 ProLeuLeuGlnArgAlaProArgAlaGluProAlaProValCysCysSerAlaArgThr 40
Db 144 CCGCTCCCTGCAGCGCGCCCCGCGGCGGAACCCGCTCCAGTATGCTGCTCGTTAC 203

QY 41 AsnLeuAlaPheLeuSerPheGlyPhePheValLeuTyrSerLeuArgValAsnLeu 60
Db 204 AACCTAGCATTTTGTCTTTTGGTTTCTGCTCTCTATTCTATTACGGGTGAATCTG 263

QY 61 SerValAlaLeuValAspMetValAspSerAsnThrThrAlaLysAspAsnArgThrSer 80
Db 264 AGCGTTGCACCTAGTGGACATGGTGATTCAAACCAACTGCGCCAAAGATAATAGAAGTCC 323

QY 81 TyrGluCysAlaGluHisSerAlaProIleLysValLeuHisAsnGlnThrGlyLysLys 100
Db 324 TACGAGTGTGAGAGCATTTGCTCCCATAAAGTTCTTCAACCAACCGGGTAAAAAG 383

QY 101 TyrArgTyrAspAlaGluThrGlnGlyTyrIleLeuGlySerPhePheTyrGlyTyrIle 120
Db 384 TACCGTGGGATGCAGAACTCAAGATGGATTCTCGGATCTTTTCTATGGGTACATC 443

QY 121 IleThrGlnIleProGlyGlyTyrValAlaSerArgSerGlyGlyLysLeuLeuGly 140
Db 444 ATCACAAATTCCTGGAGGATATGTTGCCAGCAGAGTGGGGGAGAGCTGTTGCTAGGA 503

QY 141 PheGlyIlePheAlaThrAlaIlePheThrLeuPheThrProLeuAlaAlaAspPheGly 160
Db 504 TTCGGGATCTTTGTGTACAGCTATCTTCACTCCCTGTTCACTCCCTCGCTGCAGATTCGGA 563

QY 161 ValGlyAlaLeuValAlaLeuArgAlaLeuGluGlyLeuGlyGluGlyValThrTyrPro 180
Db 564 GTCGAGGCCCTTGTTCACCTCAGGGCCTAGAGGGCTAGGAGGGGTGTCAATATCCA 623

QY 181 AlaMetHisAlaMetTyrSerSerTyrAlaProProLeuGluArgSerLysLeuSer 200
Db 624 GCCATGCATGCCATGTGTCTTCAATGGGCTCCCTCTTGAAGAAGCAAGCTTCTGAGT 683

QY 201 IleSerTyrAlaGlyAlaGlnLeuGlyThrValValSerLeuProLeuSerGlyValIle 220
Db 684 ATTTTCATATGAGGAGCACAACCTTGGGACAGTAGTTCTCTCTCTTCTTCTGGAGTAAT 743

QY 221 CysTyrTyrMetAsnTyrThrTyrValPheTyrPheGlyIleValGlyIleIleTyr 240
Db 744 TGCTACTATATGAATGGACTTATGTCTCTTATTTCTTGGCATTTGGTAATCATCTGG 803

QY 241 PheIleLeuTyrIleCysLeuValSerAspThrProGluThrHisLysThrIleThrPro 260
Db 804 TTTATTTTATGGATCTGCTTAGTTAGTGATACACAGAACTCAAGAACAATCACTCCC 863

QY 261 TyrGluLysGluTyrIleLeuSerSerLeuLysAsnGlnLeuSerSerGlnLysSerVal 280
Db 864 TATGAAGAAGGAGTATATCTTTCATCAATAAAAAATCAGCTCTCTTACAGAAGTCAAGT 923

QY 281 ProTyrIleProMetLeuLysSerLeuProLeuTyrAlaIleValAlaHisPheSer 300
Db 924 CCGTGGATACCTATGCTGAAATCACTGCCACTTTGGGCTATTGTGTCGTGCACATTTTCT 983

QY 301 TyrAsnTyrThrPheTyrThrLeuLeuThrLeuLeuProThrTyrMetLysGluValLeu 320
Db 984 TACAACCTGGACTTTTATATCTTTGTTGACCTTATTGCTACTTACATGAAGAAAGTCCTA 1043

QY 321 ArgPheAsnIleGlnGluAsnGlyPheLeuSerAlaValProTyrLeuGlyCysTyrLeu 340
Db 1044 AGGTTCAATATTCAAGAGAATGGGTTTTTATCTGAGTCCCTTATTAGGTGTGTGTTA 1103

QY 341 CysMetIleLeuSerGlyGlnAlaAlaAspAsnLeuArgAlaArgTyrAsnPheSerThr 360
Db 1104 TGTATGATCCTGTCGGCTCAAGCTGCTGACAAATTTAAGGGCAAGATGGAATTTTCAACT 1163

QY 361 LeuTyrValArgArgValPheSerLeuIleGlyMetIleGlyProAlaIlePheLeuVal 380
Db 1164 CTGTGGGTTTGCAGAGATTTTATAGCTTTATAGGGATGATGGACCTGGCATATTCTGTT 1223

QY 381 AlaAlaGlyPheIleGlyCysAspTyrSerLeuAlaValAlaPheLeuThrIleSerThr 400
Db 1224 GCCGAGATTTATAGGCTGTGATTATCTCTTGGCTGTGATTTCTTAACCATATCAACA 1283

QY 401 ThrLeuGlyGlyPheCysSerSerGlyPheSerIleAsnHisLeuAspIleAlaProSer 420
Db 1284 ACCCTGGAGGCTTTTGTCTCTCTGGATTATAGCATCAACCATCTGGACATTTGCTCTTCG 1343

QY 421 TyrAlaGlyIleLeuLeuGlyIleThrAsnThrPheAlaThrIleProGlyMetIleGly 440
Db 1344 TATGCTGGTATTTCTCTGGGCATCACAAATACCTTTGCCACTATTCTGGAATGATTTGG 1403

QY 441 ProfileAlaArgSerLeuThrProGluAsnThrIleGlyGluTyrGlnThrValPhe 460
Db 1404 CCCATCATTTGCCAAGCTTTTACCCCTGAGAACACTATTGGAGAAATGGCAACTGTTTTC 1463

QY 461 CysIleAlaAlaAlaIleAsnValPheGlyValaIlePhePheThrLeuPheAlaLysGly 480
Db 1464 TGCATCGCTGCTATCAATGATTATTGGTGCCATTTTCTTCACTATTTCGCCAAGGT 1523

QY 481 GluValGlnAsnTyrAlaIleSerAspHisGlnGlyHisArgAsn 495
Db 1524 GAAGTGCAAAACCTGGGCCATCAGTGATCAACCAAGGACACAGAAAC 1568

RESULT 2
AAD10326
ID AAD10326 standard; DNA; 2844 BP.
XX

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2005, 02:51:19 ; Search time 6301.05 Seconds
(without alignments)
3806.563 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 2617

Sequence: 1 MKPSVSLAPSGEGSDRT.....LPKGEVQNWAIHQHGRN 495

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09776865/runat_08072005_175549_23371/app_query.fasta_1.1358
-DB=GenEmbl -QFMT=fastap -SURFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09776865@cgn_1_1_8225@runat_08072005_175549_23371 -NCPUP=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2617	100.0	2844	4	AF244578	Ovis aries
2	2617	100.0	2844	6	BD248126	GBS toxin
3	2617	100.0	2844	6	AX207626	Sequence
4	2329	89.0	2512	6	AX138494	Sequence

5	2329	89.0	2512	9	HGA387747	Homo sapi
6	2329	89.0	2602	6	BD248125	GBS toxin
7	2329	89.0	2930	6	BD248129	GBS toxin
8	2329	89.0	2930	6	AX207624	Sequence
9	2329	89.0	2930	9	AF244577	Homo sapi
10	2329	89.0	3292	9	BC020961	Homo sapi
11	2329	89.0	3329	6	CQ776623	Sequence
12	2329	89.0	3362	6	CQ412026	Sequence
13	2302	88.0	2712	6	CQ783928	Sequence
14	2302	88.0	2712	6	BD127905	Primer fo
15	2302	88.0	2712	9	AK075320	Homo sapi
16	2068	79.0	1485	6	BD248130	GBS toxin
17	2068	79.0	3152	10	BC058785	Mus muscu
18	2022	77.3	3121	5	AJ719840	Gallus ga
19	1971	75.3	1485	6	BD248131	GBS toxin
20	1121	42.8	1963	3	AK114957	Ciona int
21	1104	42.2	853	6	CQ782221	Sequence
22	1104	42.2	853	6	BD126930	Primer fo
23	1081.5	41.3	2019	3	AK114801	Ciona int
24	1022.5	39.1	838	6	CQ780684	Sequence
25	1022.5	39.1	838	6	BD125393	Primer fo
26	991.5	37.9	1939	6	CQ575136	Sequence
27	985.5	37.7	1641	6	CQ596652	Sequence
28	985.5	37.7	1786	6	CQ580608	Sequence
29	985.5	37.7	1841	3	AY060776	Drosophil
30	981.5	37.5	4899	6	CQ575135	Sequence
31	981.5	37.5	3344	2	AC014246	Drosophil
32	981.5	37.5	167926	3	AC023685	Drosophil
33	981.5	37.5	174157	3	AC023711	Drosophil
34	981.5	37.5	332029	3	AE003491	Drosophil
35	981	37.5	2085	3	BT010092	Drosophil
36	980	37.4	1811	6	AR316881	Sequence
37	980	37.4	1811	6	AX711883	Sequence
38	979	37.4	3838	9	HSA459241	Homo sapi
39	979	37.4	3952	6	CQ851102	Sequence
40	979	37.4	3952	9	AK128319	Homo sapi
41	976	37.3	1767	10	AV117026	Rattus no
42	976	37.3	3172	10	RN0491795	Rattus no
43	969.5	37.0	1856	10	AF510321	Mus muscu
44	969.5	37.0	4481	10	BC042593	Mus muscu
45	969	37.0	2528	6	AX709538	Sequence

ALIGNMENTS

RESULT 1
AF244578
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AF244578
Ovis aries membrane glycoprotein SP55 (sp55) mRNA, complete cds.
AF244578.1 GI:9719375
Ovis aries (sheep)
Ovis aries
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Caprinae; Ovis.
1 (bases 1 to 2844)
2 (bases 1 to 2844)
Fu, C., Bardhan, S., Cetateanu, N.D., Lloyd, S.R., Yan, H., P.,
Carter, C.E., Shi, E., Venkov, C., Yakes, M.F., Page, D.L. and H.C.G.
Identification of a novel membrane protein from mammalian cells
that interact with the anti-pathogenic compound CM101
Unpublished
2 (bases 1 to 2844)
Fu, C., Bardhan, S., Cetateanu, N.D., Lloyd, S.R. and Heltqvist, C.G.
Direct Submission
Submitted (13-MAR-2000) Biochemistry, Vanderbilt University, School
of Medicine, 23rdPierce, Nashville, TN 37232-0146, USA

Location/Qualifiers
1. .2844
/organism="Ovis aries"
/mol_type="mRNA"
/db_xref="taxon:9940"

/cell_type="endothelial cell"
 /tissue_type="lung"
 1..2844
 /gene="sp55"
 84..1571
 /gene="sp55"
 /note="similar to Na+ inorganic phosphate transporter"
 /codon_start=1
 /product="membrane glycoprotein SP55"
 /protein_id="AAF97770.1"
 /db_xref="GI:9719376"
 /translation="MKSVPDLAPSDDEGSDRPLQRPAPBPVCCSARNLAF
 LSPFGFVLYSLRNLVALVDMSNTAKDNRTSYECAHSAPIKVLHNQTKKVR
 WDAETQWILGSPFYGYITQIPGVYASRGSKLLGFGIFATAIFLFTPLAADF
 VGLVALRALGELGEGVTPAMHAMWSWAPLERSKLLSISYAGAOLGTVVSLPLSG
 VICVYMNVTYVYFPGIVGIIWFLILWICLVSDTPETHKTIITPYEKVITLSSKNOLSS
 QKSPVIMPLKSLPLMALIVAHFSYNWTFYILLILLPYMKVEVERFNPQENGFLSAPV
 YLGCWMLSGADNRNFSTLMWRVRFSLIGMIGPAIFLVAAGFTIGCDYSIA
 VAFITISTLGGFCSSGSIINHLDIAPSYAGILGINTFATIFPMIGPIIARSLTPE
 NTIGEWQTVFCIAAALNVFGAIFFTLPAKGEVONWAIISDHQHRN"

ORIGIN

Alignment Scores:

Pred. No.: 3 528-226 Length: 2844
 Score: 2617.00 Matches: 495
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-776-865-4 (1-495) x AF244578 (1-2844)

QY 1 MetLysSerProValSerAspLeuAlaProSerAspGlyGluGlySerAspArgThr 20
 DB 84 ATGAAGTCCCGGTTTCGGACTTACCCCGAGCGAGCGAGGGCTCGACCGACA 143
 QY 21 ProLeuLeuGlnArgAlaProArgAlaGluProAlaProValCysCysSerAlaArgTyr 40
 DB 144 CCGCTCCTGCAGCGCGCCCGCGCGGACCCGCTCCAGTATGCTGCTGCTGTTAC 203
 QY 41 AnLeuAlaPheLeuSerPhePheGlyPhePheValLeuTyrSerLeuArgValAsnLeu 60
 DB 204 AACCTAGCATTTTTCCTCTTTTGGTTTCTTCGTTCTCTATTCTATTACGGGTGAATCTG 263
 QY 61 SerValAlaLeuValAspMetValAspSerAsnThrThrAlaLysAspAsnArgThrSer 80
 DB 264 AGCGTTGACACTAGTGACATGCTGGATTCAACACTGCCAAGATGAATAGACGTCC 323
 QY 81 TyrGluCysAlaGluHisSerAlaProIleLysValLeuHisAsnGlnThrGlyLysLys 100
 DB 324 TACGAGTGTGCAGAGCATTCGCTCCCATAAAAGTTCTTCAACCAACCGGTAAAAAG 383
 QY 101 TyrArgTyrAspAlaGluThrGlnGlyTyrIleLeuGlySerPhePheTyrGlyTyrIle 120
 DB 384 TACCGGTGGGATGCAGAACTCAAGGATGGATTCCTCGGATCTTTTTTCTATGGCTACATC 443
 QY 121 IleThrGlnIleProGlyGlyTyrValAlaSerArgSerGlyGlyLysLeuLeuGly 140
 DB 444 ATCACAAATTCCTGGAGGATATGTTGCCAGCAGAGTGGGGGAAGCTGTGCTAGGA 503
 QY 141 PheGlyIlePheAlaThrAlaIlePheThrLeuPheThrProLeuAlaAlaAspPheGly 160
 DB 504 TTCGGGATCTTGTACAGCTATCTTCACCTGTTCACTCCCTCGCTGCAGATTTCGGA 563
 QY 161 ValGlyAlaLeuValAlaLeuArgAlaLeuGluGlyLeuGlyGluGlyValThrTyrPro 180
 DB 564 CTCGGAGCCCTTGTTCACCTCAGGCACTAGAAAGGGCTAGGAGAGGGTGTACATATCCA 623
 QY 181 AlaMetHisAlaMetTyrSerSerTyrAlaProProLeuGluArgSerLysLeuLeuSer 200
 DB 624 GCCATGCAATGCCATGGTGTTCATGGGCTCCCTCTTGAAGAAGCAAGCTTCTAGT 683
 QY 201 IleSerTyrAlaGlyAlaGlnLeuGlyThrValValSerLeuProLeuSerGlyValIle 220

DB 684 ATTTTCATATGCAGGAGCACAACTTGGGACAGTAGTTTCTCTCTCTCTTCTGGAGTAAAT 743
 QY 221 CysTyrTyrMetAsnTyrThrTyrValPheTyrPhePheGlyIleValGlyIleIleTyr 240
 DB 744 TGCCTACTATATGAATGGACTTATGTCTCTATTCTTTTGGCATTTGGTAATCATCTGG 803
 QY 241 PheIleLeuTyrIleCysLysLeuValSerAspThrProGluThrHisLysThrIleThrPro 260
 DB 804 TTTATTTATGGAICTGCTTAGTTAGTGATACACACAGAACTCACAGACAATCACTCCC 863
 QY 261 TyrGluLysGluTyrIleLeuSerSerLysLysAsnGlnLeuSerSerGlnLysSerVal 280
 DB 864 TATCAAAAGGAGTAGTATATTTCTTTCATCAATAAAAAATCAGCTCTCTTACAGAACTCAGT 923
 QY 281 ProTyrIleProMetLeuLysSerLeuProLeuTyrAlaIleValAlaIleHisPheSer 300
 DB 924 CCGTGGATACCTATGCTGGAATCACTGCCACTTTGGGCTATTGTGCTTGCAATTTTTCT 983
 QY 301 TyrAsnTyrThrPheTyrThrLeuLeuLeuProThrTyrMetLysGluValLeu 320
 DB 984 TACAACCTGGACTTTTATACCTTTGTCACCTTATTGCTACTTACATGAAGAAAGTCTTA 1043
 QY 321 ArgPheAsnIleGlnGluAsnGlyPheLeuSerAlaValProTyrLeuGlyCysTyrLeu 340
 DB 1044 AGGTTCAATATTCAGAGAAATGGGTTTTTATCTCAGTCCCTTATTAGGTGTGTGTTA 1103
 QY 341 CysMetIleLeuSerGlyGlnAlaAspAsnLeuArgAlaArgTyrAsnPheSerThr 360
 DB 1104 TGTATGATCCCTGTCGGGTCAAGCTGCTGACAAATTAAGGGCAAGTGAATTTTTCAACT 1163
 QY 361 LeuTyrValArgArgValPheSerLeuIleGlyMetIleGlyProAlaIlePheLeuVal 380
 DB 1164 CTGTGGGTTCGAAGAGTTTTTAGCTTTATAGGGATGATGGACCTGGCATATTCTGTT 1223
 QY 381 AlaAlaGlyPheIleGlyCysAspTyrSerLeuAlaValAlaPheLeuThrIleSerThr 400
 DB 1224 GCCGACGAGATTTATAGGCTGTGATTATTCCTTGGCTGTGCAATTCCTAAACCATATCAACA 1283
 QY 401 ThrLeuGlyGlyPheCysSerSerGlyPheSerIleAsnHisLeuAspIleAlaProSer 420
 DB 1284 ACCCTGGAGGCTTTTGTCTTCTGGATTTAGCATCAACCATCTGGACATTCGCTCTTCG 1343
 QY 421 TyrAlaGlyIleLeuLeuGlyIleThrAsnThrPheAlaThrIleProGlyMetIleGly 440
 DB 1344 TATGCTGGTATTCCTCTGGGATCACAATATACCTTTGCCACTATTCCTGGAATCATTTGG 1403
 QY 441 ProIleIleAlaArgSerLeuThrProGluAsnThrIleGlyGluTyrGlnThrValPhe 460
 DB 1404 CCCATCATTTGCCAGAGTCTTACCCTTGAGAACACTATTGGAGAATGGCAACTGTTTTC 1463
 QY 461 CysIleAlaAlaIleAsnValPheGlyValIlePhePheThrLeuPheAlaLysGly 480
 DB 1464 TGCATCGCTGCTGTATCAATGATTTGGTGCCATTTTCTTTCACACTATTTCGCCAAAGGT 1523
 QY 481 GluValGlnAsnTyrAlaIleSerAspHisGlnGlyHisArgAsn 495
 DB 1524 GAAATGCBAACCTGGGCCATCAGTGATCACCAGGACACAGAAC 1568
 RESULT 2
 BD248126
 LOCUS
 DEFINITION
 GBS toxin receptor.
 BD248126
 ACCESSION
 BD248126.1 GI:33057896
 VERSION
 JP 2002524027-A/2.
 KEYWORDS
 Ovis sp.
 SOURCE
 Ovis sp.
 ORGANISM
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Caprinae; Ovis.
 REFERENCE
 1 (bases 1 to 2844)
 AUTHORS
 Hellerqvist,C.G. and Fu,C.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2005, 03:19:49 ; Search time 4969.57 Seconds
(without alignments)
4105.478 Million cell updates/sec

Title: US-09-776-865-2

Perfect score: 2836

Sequence: 1 MAAGAMTTPRPVQPARPGSG.....LPAKGEVQNWALNDHGHHRH 536

Scoring table: BLOSUM62

-Q/cgn2 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+p2n.model -DB=xlh
-Q/cgn2 10.0 , Xgapext 0.5
-Ygapop 10.0 , Ygapext 0.5
-Fgapop 6.0 , Fgapext 7.0
-Delop 6.0 , Delext 7.0
-DOALIGN=200 -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09776865@cgn 1.1.6461@runat.08072005 175549 23379 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2752	97.0	3189	3 CR618872	CR618872 full-length
2	2312	81.5	3202	3 AK029102	AK029102 Mus muscu
3	1772	62.5	1051	1 AL550137	AL550137 AL550137
4	1699.5	59.9	1059	5 BX425026	BX425026 BX425026
5	1651	58.2	1038	5 BX439809	BX439809 BX439809
6	1347	47.5	947	7 COS79484	COS79484 ILLUMIGEN
7	1236	43.6	754	4 BI907284	BI907284 603063858
8	1193	41.7	737	7 CK833737	CK833737 4057771 B
9	1174.5	41.4	901	6 CD106410	CD106410 AGENCOURT

10	1170	41.3	721	7	CK833565	CK833565 4057467 B
11	1166	41.1	708	7	CR763802	CR763802 DKF204690
12	1137	40.1	773	1	AJ734267	AJ734267 AJ734267
13	1129	39.8	730	6	CA327442	CA327442 UI-M-FYO-
14	1120	39.5	740	6	CB317739	CB317739 AGENCOURT
15	1115	39.3	801	4	BI661062	BI661062 603304362
16	1101.5	38.8	663	7	CV023522	CV023522 483 Full
17	1075	37.9	813	7	CF618610	CF618610 AGENCOURT
18	1062	37.4	676	2	BE867611	BE867611 601443127
19	1057.5	37.3	791	7	CV110989	CV110989 AGENCOURT
20	1056	37.2	913	5	BX348297	BX348297 BX348297
21	1051.5	37.1	831	3	AK087395	AK087395 Mus muscu
22	1051	37.1	605	5	BX479639	BX479639 DKF20686F
23	1041.5	36.7	736	5	BP223105	BP223105 BP223105
24	1024.5	36.1	736	4	BI860521	BI860521 603386787
25	1015.5	35.8	770	7	CF289359	CF289359 AGENCOURT
26	1015	35.8	582	5	BP302998	BP302998 BP302998
27	1014	35.8	584	5	BP288606	BP288606 BP288606
28	1010	35.6	581	5	BP298092	BP298092 BP298092
29	1008	35.5	1007	4	BM804862	BM804862 AGENCOURT
30	1003	35.4	582	5	BP287958	BP287958 BP287958
31	1002	35.3	583	4	BM838178	BM838178 K-EST0114
32	1001	35.3	581	5	BP297030	BP297030 BP297030
33	995	35.1	579	5	BP344739	BP344739 BP344739
34	995	35.1	602	5	BP349262	BP349262 BP349262
35	994	35.0	581	5	BP281761	BP281761 BP281761
36	991	34.9	570	5	BP274537	BP274537 BP274537
37	991	34.9	626	2	BF676817	BF676817 602084380
38	991	34.9	711	4	BI697765	BI697765 603346858
39	990	34.9	580	5	BP285113	BP285113 BP285113
40	985	34.7	582	5	BP287215	BP287215 BP287215
41	979	34.5	582	5	BP367879	BP367879 BP367879
42	978	34.5	616	5	BP238694	BP238694 BP238694
43	975	34.4	581	5	BP285186	BP285186 BP285186
44	975	34.4	582	5	BP288187	BP288187 BP288187
45	973	34.3	630	6	CB138761	CB138761 K-EST0191

ALIGNMENTS

CR618872 3189 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DI040YK17 of Placenta Cot 25-normalized
of Homo sapiens (human).

CR618872

CR618872.1 GI:50499679

HTC; CDS; CDS; CDS

HTC; CDS; CDS; CDS

Homo sapiens (human)

Homo sapiens

Homo sapiens

1 (bases 1 to 3189)

Li W.B., Gruber C., Jessee J. and Polayes D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Feng Liang Email: fliang@lifetech.com URL:

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue

Genoscope.

2 (bases 1 to 3189)

Direct Submission

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage:

BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr

- Web: www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

Location/Qualifiers

1..3189

/organism="Homo sapiens"

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1040YK17"
 /tissue_type="Placenta Cot 25-normalized"
 /plasmid="pCWSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.: 1.22e-270 Length: 3189
 Score: 2752.00 Matches: 520
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 97.04% Indels: 0
 DB: 3 Gaps: 0

US-09-776-865-2 (1-536) x CR618872 (1-3189)

QY 17 ProGlyGlyPheGlyLeuSerGlyArgArgSerLeuLeuGlyGlnValAlaSerThrPro 36
 Db 1 CCGGGGGGCTCGGGCTGTGGGGCGGGCTCCCTTCTGCGCAGGTGGCAGTACACCT 60
 QY 37 AlaHisValGlyValMetArgSerProValArgAspLeuAlaArgAsnAspGlyGluGlu 56
 Db 61 GCTCAGTAGGGCTCATGAGGCTCCGGTTCGAGACCTGGCCCGAAGCATGGCGAGGAG 120
 QY 57 SerThrAspArgThrProLeuLeuProGlyAlaProArgAlaGluAlaAalaProValCys 76
 Db 121 AGCAGCGGCGCACGCCCTCTTACCGGGCGGCCCCACCGGGCCGAGCCGCTCCAGTGTGC 180
 QY 77 CysSerAlaArgTyAsnLeuAlaAlaPheGlyPheGlyPhePheIleValTyAla 96
 Db 181 TGCTCTGCTCGTACAACTTAGCAATTTTGGGCTTTTGGTCTTCTTCTTCTTCTTCTTCTT 240
 QY 97 LeuArgValAsnLeuSerValAlaLeuValAspMetValAspSerAsnThrThrLeuGlu 116
 Db 241 TTACGTGTGAATCTCAGGTGTTCGGTGTAGGTATGGTATGTTTCAATACAACTTTAGAA 300
 QY 117 AspAsnArgThrSerLysAlaCysProGluHisSerAlaProIleLysValHisAsn 136
 Db 301 GATAATAGAACTTCCAGGGCGTTCAGAGCACTTCTGCTCCCAATAAAGTTTCATCATAT 360
 QY 137 GlnThrGlyLysLysTyGlnThrAspAlaGluThrGlnGlyThrIleLeuGlySerPhe 156
 Db 361 CAAACGGGTAAAGATGCAATGGAATGGATGAGAACTCAAGGATGATTCGCGTTCCTTT 420
 QY 157 PheTyGlyTyIleIleThrGlnIleProGlyGlyTyValAlaSerLysIleGlyGly 176
 Db 421 TTTTATGCTACATCATCACAGATTTCTGAGGATATGTTGCCAGCAAAATAGGGGG 480
 QY 177 LysMetLeuLeuGlyPheGlyIleLeuGlyThrAlaValLeuThrLeuPheThrProIle 196
 Db 481 AAAATGCTGTAGGATTTGGGATCTTGGCACTGCTGCTCCACCTGTCTCACTCCCAT 540
 QY 197 AlaAlaAspLeuGlyValGlyProLeuIleValLeuArgAlaLeuGlyGlyLeuGlyGlu 216
 Db 541 GCTGCAGATTTAGGAGTTGGACCACTCATTTGACTCAGAGCACTAGAGGACTAGGAG 600
 QY 217 GlyValThrPheProAlaMetHisAlaMetTrpSerSerTrpAlaProLeuGluArg 236
 Db 601 GGTGTATCATTTCCAGCATGATGCCATGTGTCTTCTTGGGCTCCCGCTTGAAGA 660
 QY 237 SerLysLeuLeuSerIleSerTyAlaGlyAlaGlnLeuGlyThrValIleSerLeuPro 256
 Db 661 AGCAAACTTCTTAGCATTTTCATATGCAGGAGCACAGCTTGGACAGTAAATTTCTTCT 720
 QY 257 LeuSerGlyIleIleCysTyTrpMetAsnTrpThrTyValPheTyPhePheGlyThr 276
 Db 721 CTTTCTGGAATAATTTGCTACTATATGAATTTGACTTATGCTTCTTCTTCTTCTTCTTCT 780
 QY 277 IleGlyIlePheThrPheLeuLeuTrpIleTrpLeuValSerAspThrProGlnLysHis 296
 Db 781 ATTGGAATATTTGGTCTTTTGTGGATCTGGTATGTAGTATGTAGTACACCAACCAAAACAC 840

QY 297 LysArgIleSerHisTyTrpGlyGlyGluTyTrpIleLeuSerSerLeuArgAsnGlnLeuSer 316
 Db 841 AAGAGAATTTCCCATTTAAGAAAGGAATACATCTCTTTCATCATTAAGAAATCAGCTTCT 900
 QY 317 SerGlnLysSerValProTrpValProIleLeuLysSerLeuProLeuTrpAlaIleVal 336
 Db 901 TCACAGAAAGTCAGTCCGCGGTACCCATTTTAAATCCCTGGCCACTTTGGGCTATCGTA 960
 QY 337 ValAlaHisPheSerTyAsnTrpThrPheTyTrpLeuLeuThrLeuLeuProThrTyTrp 356
 Db 961 GTTGCACACTTTTCTTACAACTGGACTTTTATATCTTATTTGACATATTTGCCACTTAT 1020
 QY 357 MetLysGluIleLeuArgPheAsnValGlnGluAsnGlyPheLeuSerSerLeuProTyTrp 376
 Db 1021 ATGAAGGAGATCCTAAGGTTCAATGTTCAAGAGAATGGGTTTTATCTTCATCTGCTTAT 1080
 QY 377 LeuGlySerTrpLeuCysMetIleLeuSerGlyGlnAlaAlaAspAsnLeuArgAlaLys 396
 Db 1081 TTAGGCTCTTGGTTATGTATGATCTCTGTGGTCAAGCTGTGCAAAATTTAAGGGCAAAA 1140
 QY 397 TrpAsnPheSerThrLeuCysValArgArgIlePheSerLeuIleGlyMetIleGlyPro 416
 Db 1141 TGGAAATTTTCACTTTATGTGTCGAGAAATTTTACGCTTATAGCAATGATTTGGACT 1200
 QY 417 AlaValPheLeuValAlaAlaGlyPheIleGlyCysAspTyTrpSerLeuAlaValAlaPhe 436
 Db 1201 GCAGTATTTCTGGTAGCTGTGGCTTCATTTGGCTGTGATTTCTTTGGCGGTTGCTTTC 1260
 QY 437 LeuThrIleSerThrThrLeuGlyGlyPheCysSerSerGlyPheSerIleAsnHisLeu 456
 Db 1261 CTAACTATATCAACAACACTGGGAGGCTTTGCTCTCTGGAATTAGCATCAACCATCTG 1320
 QY 457 AspIleAlaProSerTyTrpAlaGlyIleLeuLeuGlyIleThrAsnThrPheAlaThrIle 476
 Db 1321 GATATTGCTCTTGTATGCTGGTATCTCTCTGGGCATCAACAATACATTTGCCACTATT 1380
 QY 477 ProGlyMetValGlyProValIleAlaLysSerLeuThrProAspAsnThrValGlyGlu 496
 Db 1381 CCAGGAATGGTTGGCGCGCTCATTTGCTAAAGTCTGACCCCTGATAACACTGTTGGAGAA 1440
 QY 497 TrpGlnThrValPheTyTrpIleAlaAlaIleValPheGlyAlaIlePhePheThr 516
 Db 1441 TGGCAAAACCGTGTCTATATTGCTGCTGCTATTAATGTTTGGTGCCATTTCTTTTACA 1500
 QY 517 LeuPheAlaLysGlyGluValGlnAsnTrpAlaLeuAsnAspHisHisGlyHisArgHis 536
 Db 1501 CTATTCGCAAAAGTGAAGTACAAAACCTGGGCTCTCAATGATCACCATGGACACAGAC 1560
 RESULT 2
 AK029102 3202 bp mRNA linear HTC 03-APR-2004
 LOCUS Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
 DEFINITION library, clone:4732491M05 product:SIALIN homolog (Homo sapiens),
 full insert sequence.
 ACCESSION AK029102
 VERSION AK029102.1 GI:26325087
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Mech. Enzymol. 303, 19-44 (1999)
 99279253
 PUBMED 10349636
 REFERENCE 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2005, 07:16:14 ; Search time 910.316 Seconds
(without alignments)
3636.712 Million cell updates/sec

Title: US-09-776-865-2

Perfect score: 2836

Sequence: 1 MAAGAMTPRPVQPARPGSF.....LFKAGEVQNALNDHGHRRH 536

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6330943 seqs, 3139157217 residues

Total number of hits satisfying chosen parameters: 12661886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spoel/US09776865/runat 08072005 175551 23472/app query.fasta 1.1358
-DB=Published Applications NA -OPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLAL=IGNF=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09776865 @CNC 1.1 1065 @runat 08072005 175551 23472
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US05_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2836	100.0	2930	9	US-09-776-865-1	Sequence 1, Appli
2	2836	100.0	2930	21	US-10-823-506-7	Sequence 7, Appli
3	2804	98.9	2626	14	US-10-198-846-12482	Sequence 12482, A
4	2804	98.9	3362	10	US-09-814-353-19097	Sequence 19097, A
5	2770.5	97.7	3329	19	US-10-755-889-587	Sequence 587, App
6	2761	97.4	3292	21	US-10-887-553A-858	Sequence 858, App
7	2717	95.8	2602	21	US-10-823-506-1	Sequence 1, Appli
8	2614	92.2	1488	13	US-10-098-841-322	Sequence 322, App
9	2338	82.4	2844	9	US-09-776-865-3	Sequence 3, Appli
10	2338	82.4	2844	21	US-10-823-506-3	Sequence 3, Appli
11	2217	78.2	1485	21	US-10-823-506-9	Sequence 9, Appli
12	1967	69.4	1485	21	US-10-823-506-11	Sequence 11, Appli
13	1521	53.6	1872	17	US-10-264-049-302	Sequence 302, App
14	1515	53.4	1651	17	US-10-264-237-946	Sequence 946, App
15	973	34.3	3982	18	US-10-205-331-8	Sequence 8, Appli
16	973	34.3	3982	20	US-10-734-731-11	Sequence 11, Appli
17	973	34.3	3982	20	US-10-807-500-11	Sequence 11, Appli
18	971	34.2	2528	20	US-10-734-731-13	Sequence 13, Appli
19	971	34.2	2528	20	US-10-807-500-13	Sequence 13, Appli
20	968	34.1	3946	20	US-10-734-731-9	Sequence 9, Appli
21	968	34.1	3946	20	US-10-807-500-9	Sequence 9, Appli
22	968	34.1	3946	21	US-10-887-553A-1016	Sequence 1016, Ap
23	962	33.9	3423	16	US-10-233-045-21	Sequence 21, Appli
24	956	33.7	1811	9	US-09-740-041-1	Sequence 1, Appli
25	956	33.7	1811	16	US-10-389-967-1	Sequence 1, Appli
26	955	33.7	1767	22	US-10-499-731-29	Sequence 29, Appli
27	955	33.7	1850	22	US-10-499-731-45	Sequence 45, Appli
28	955	33.7	3838	21	US-10-887-553A-1200	Sequence 1200, Ap
29	945	33.3	1803	22	US-10-499-731-47	Sequence 47, Appli
30	945	33.3	1822	22	US-10-499-731-58	Sequence 58, Appli
31	943	33.3	2716	17	US-10-314-790-6	GENERAL INFORMA
32	942.5	33.2	2366	20	US-10-734-731-1	Sequence 1, Appli
33	942.5	33.2	2366	20	US-10-807-500-1	Sequence 1, Appli
34	942	33.2	2716	20	US-10-734-731-3	Sequence 3, Appli
35	942	33.2	2716	20	US-10-807-500-3	Sequence 3, Appli
36	940.5	33.2	1701	18	US-10-287-226-299	Sequence 299, App
37	940.5	33.2	1975	18	US-10-296-115-461	Sequence 461, App
38	939	33.1	2024	18	US-10-152-319A-2039	Sequence 2039, Ap
39	939	33.1	2024	20	US-10-734-731-5	Sequence 5, Appli
40	939	33.1	2024	20	US-10-807-500-5	Sequence 5, Appli
41	936	33.0	2836	20	US-10-734-731-7	Sequence 7, Appli
42	936	33.0	2836	20	US-10-807-500-7	Sequence 7, Appli
43	935	33.0	2607	9	US-09-915-181A-1	Sequence 1, Appli
44	852	30.0	494	10	US-09-814-353-12829	Sequence 12829, A
45	837	29.5	498	10	US-09-814-353-46	Sequence 46, Appli

ALIGNMENTS

RESULT 1
US-09-776-865-1
; Sequence 1, Application US/09776865
; Patent No. US20020061846A1
; GENERAL INFORMATION:
; APPLICANT: Hellergvist, Carl
; TITLE OF INVENTION: Methods for Preventing or Attenuating Pathoangiogenic Conditions
; FILE REFERENCE: 22100-0100 46126-252687
; CURRENT APPLICATION NUMBER: US/09/776,865
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 60/179,870
; PRIOR FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2930

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (263)..(1870)
US-09-776-865-1

Alignment Scores:
Pred. No.: 2,376-282 Length: 2930
Score: 2836.00 Matches: 536
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-776-865-2 (1-536) x US-09-776-865-1 (1-2930)

Qy 1 MetAlaAlaGlyAlaMetThrProArgProValGlnProAlaArgProGlyGlyPhe 20
Db 263 ATGGCGGGGGGGCGATGACACCGCCCGCGGGTCCAGCCAGCTCGGCCGGGGCTTC 322

Qy 21 GlyLeuSerGlyArgSerLeuLeuCysGlnValAlaSerThrProAlaHisValGly 40
Db 323 GGGCTGTGGGCGGGCTCCCTCTCTCCAGGTGGCGAGTACACCTGCTCACGTAGGC 382

Qy 41 ValMetArgSerProValArgAspLeuAlaArgAsnAspGlyGluGluSerThrAspArg 60
Db 383 GTCATGAGGTCTCCGGTTCGAGACCTGGCCCGGAACGATGGCGAGGAGCAGCAGCCGC 442

Qy 61 ThrProLeuLeuProGlyAlaProArgAlaGluAlaAlaProValCysCysSerAlaArg 80
Db 443 ACGCTCTTTCACCGGGGGCGCCCGCCAGCGCCGAGCGCTCCAGTGTGCTCTGCTCGT 502

Qy 81 TyrAsnLeuAlaLeuAlaPhePheGlyPhePheIleValTyrAlaLeuArgValAsn 100
Db 503 TACAACCTTAGCNAATTTGGCCCTTTTGGCTTTCTTCATGTGTATGTCATTACGTGTGAT 562

Qy 101 LeuSerValAlaLeuValAspMetValAspSerAsnThrThrLeuGluAspAsnArgThr 120
Db 563 CTGAGTGTGGCTTAGTGGATATGGTAGATTCAATACAACTTTAGAAAGATAATAGAACT 622

Qy 121 SerIysAlaCysProGluHisSerAlaProIleIysValHisAsnGlnThrGlyLys 140
Db 623 TCCAAAGGGCTGCCAGAGCATTTGCTCCCAATAAAGTTTCATCAATAACAAACGGGTAA 682

Qy 141 LysTyrGlnTyrAspAlaGluThrGlnGlyTyrIleLeuGlySerPhePheTyrGlyTyr 160
Db 683 AAGTACCAATGGGATGCGAAGAACTCAAGATGGATTCTCGGTTCTTTTATGGCTAC 742

Qy 161 IleIleThrGlnIleProGlyGlyTyrValAlaSerLysIleGlyGlyLeuMetLeuLeu 180
Db 743 ATCATCACAGATTCCTGGAGGATATGTTGCCAGCAAAATAGGGGGGAAAAATGCTGCTA 802

Qy 181 GlyPheGlyIleLeuGlyThrAlaValLeuThrLeuPheThrProIleAlaAlaAspLeu 200
Db 803 GGATTTGGGATTCCTGGACACTGCTGCTCCACCCCTGTTTCATCTCCCATTTGCTGCAGATT 862

Qy 201 GlyValGlyProLeuIleValLeuArgAlaLeuGluGlyLeuGlyGluGlyValThrPhe 220
Db 863 GGAGTTGGACCACTCATTTGTACTCAGACACTAGNAGGACTAGNAGGGGTTCACATTT 922

Qy 221 ProAlaMetHisAlaMetTrpSerSerTrpAlaProProLeuGluArgSerLysLeuLeu 240
Db 923 CCAGCATGCATGCCATGTGGTCTCTTGGGCTCCCTCTTGAAGAAGCAAACTTCTT 982

Qy 241 SerIleSerTyrAlaGlyAlaGlnLeuGlyThrValIleSerLeuProLeuSerGlyIle 260
Db 983 AGCATTTGATGCGAGGAGCAGCTGGGACAGTAATTTCTTCTCTTCTTCTGGAATA 1042

Qy 261 IleCysTyrTyrMetAsnTrpThrTyrValPheTyrPheGlyThrIleGlyIlePhe 280
Db 1043 ATTGTACTATATGAATGGACTTATGCTCTTACTTTTTTGGTACTATTGGATATTT 1102

```

```

Qy 281 TrpPheLeuLeuTrpIleTTrpLeuValSerAspThrProGlnLysHisLysArgIleSer 300
Db 1103 TGGTTTCTTTTGTGGATCTGGTTAGTGTACACACACACAAAAACACACAGAGAATTTCC 1162

Qy 301 HisTyrGluLysGluTyrIleLeuSerSerLeuArgAsnGlnLeuSerSerGlnLysSer 320
Db 1163 CATTATGAAAGAGGAATACATTTCTTTTCATCATTAAGAATAGCTTTCTTTCAGAGTCA 1222

Qy 321 ValProTrpValProIleLeuLysSerLeuProLeuTrpAlaIleValValAlaHisPhe 340
Db 1223 GTGCGGTGGGTACCAATTTAAATCCCTGCCATTTGGGCTATCGTAGTTGCACATTT 1282

Qy 341 SerTyrAsnTrpThrPheTyrThrLeuLeuThrLeuLeuProThrTyrMetLysGluIle 360
Db 1283 TCTTACAACTGGACTTTTATACATTTATGACATTTATGCTACTTATATATGAGGAGATC 1342

Qy 361 LeuArgPheAsnValGlnGluAsnGlyPheLeuSerSerLeuProTyrLeuGlySerTrp 380
Db 1343 CTAAGGTTCATGTTCAGAGAAATGGGTATTTATCTTCATTGCTTATTTAGGCTCTTGG 1402

Qy 381 LeuCysMetIleLeuSerGlyGlnAlaAlaAspAsnLeuArgAlaLysTrpAsnPheSer 400
Db 1403 TTATGTATGATCCTGTCTGGTCAAGCTGCTGACAAATTTAAGGGCAAAATGGAAATTTTCA 1462

Qy 401 ThrLeuCysValArgArgIlePheSerLeuIleGlyMetIleGlyProAlaValPheLeu 420
Db 1463 ACTTTATGTGTTCGAGAAATTTTATAGCTTTATAGAAATGATTTGACCTGCGAGTATTCCTG 1522

Qy 421 ValAlaAlaGlyPheIleGlyCysAspTyrSerLeuAlaValAlaPheLeuThrIleSer 440
Db 1523 GTAGCTGCTGCTTCATTTGGCTGTGATTTCTTTGGCCGTTGCTTTCTTAACTATATATCA 1582

Qy 441 ThrThrLeuGlyGlyPheCysSerSerGlyPheSerIleAsnHisLeuAspIleAlaPro 460
Db 1583 ACAACACTGGGAGGCTTTTGTCTCTCTGGATTTAGCATCAACCATCTGGATATTGCTCCT 1642

Qy 461 SerTyrAlaGlyIleLeuLeuGlyIleThrAsnThrPheAlaThrIleProGlyMetVal 480
Db 1643 TCGTATGCTGTATCTCTCTGGCATCACAAATACATTTGCGCACTATTTCAGGAATGGTT 1702

Qy 481 GlyProValIleAlaLysSerLeuThrProAspAsnThrValGlyGluTyrGlnThrVal 500
Db 1703 GGGCCCTCATTTGTATAAGTCTGACCCCTGATAACTGTTGGAGAAATGGCAACCGTG 1762

Qy 501 PheTyrIleAlaAlaAlaIleAsnValPheGlyAlaIlePhePheThrLeuPheAlaLys 520
Db 1763 TTCTATATGCTGCTGCTATTAATGTTTGGTGCCATTTCTTTACACTATTTCGCCAA 1822

Qy 521 GlyGluValGlnAsnTrpAlaLeuAsnAspHisHisGlyHisArgHis 536
Db 1823 GGTGAAGTACAAAACCTGGGCTCTCAATGATCATCCATGAGACACACAGACAC 1870

RESULT 2
US-10-823-506-7
; Sequence 7, Application US/10823506
; Publication No. US20050002931A1
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/0105
; CURRENT APPLICATION NUMBER: US/10/823,506
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: US/09/359,167
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60-693,843
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2005, 06:18:14 ; Search time 253.183 Seconds

(without alignments)
3464.068 Million cell updates/sec

Title: US-09-776-865-2

Perfect score: 2836

Sequence: 1 MAAGAMTPRPVQPARPGF.....LFAKGEVQNALNDHGHGRH 536

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09776865/runat_08072005_175549_23397/app_query.fasta_1.1358
-DB=Issued Patents NA -QMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09776865@cgn_1_1_277@runat_08072005_175549_23397 -NCPUS=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2836	100.0	2930	4	US-09-359-167-7	Sequence 7, Appli
2	2804.5	98.9	2513	4	US-09-949-016-1834	Sequence 1834, Ap
3	2717	95.8	2602	4	US-09-359-167-1	Sequence 1, Appli
4	2338	82.4	2844	4	US-09-359-167-3	Sequence 3, Appli
5	2217	78.2	1485	4	US-09-359-167-9	Sequence 9, Appli
6	1967	69.4	1485	4	US-09-359-167-11	Sequence 11, Appli
7	956	33.7	1811	4	US-09-740-041-1	Sequence 1, Appli
8	942.5	33.2	2366	4	US-09-949-016-5483	Sequence 5483, Ap
9	942	33.2	2716	1	US-08-647-484-1	Sequence 1, Appli
10	942	33.2	2716	1	US-08-647-484-3	Sequence 3, Appli
11	942	33.2	2716	1	US-08-647-481-1	Sequence 1, Appli
12	942	33.2	2716	1	US-08-647-481-3	Sequence 3, Appli

13	942	33.2	2716	1	US-08-430-033A-1	Sequence 1, Appli
14	942	33.2	2716	1	US-08-430-033A-3	Sequence 3, Appli
15	942	33.2	2716	5	PCT-US96-05792-1	Sequence 1, Appli
16	942	33.2	2716	5	PCT-US96-05792-3	Sequence 3, Appli
17	935	33.0	2607	4	US-09-915-181A-1	Sequence 1, Appli
18	876	30.9	2626	4	US-09-949-016-745	Sequence 745, App
19	810	28.6	2612	4	US-09-270-767-13970	Sequence 13970, A
20	795	28.0	1229	4	US-08-023-655-660	Sequence 660, App
21	776	27.4	2266	2	US-08-724-394A-18	Sequence 18, Appl
22	776	27.4	2270	4	US-09-949-016-5577	Sequence 5577, Ap
23	629	22.2	1643	2	US-08-805-118-2	Sequence 2, Appli
24	629	22.2	1643	3	US-09-391-958-2	Sequence 2, Appli
25	625.5	22.1	1780	2	US-08-724-394A-19	Sequence 19, Appl
26	622	21.9	2296	4	US-09-949-016-1182	Sequence 1182, Ap
27	470.5	16.6	1699	4	US-09-739-451-815	Sequence 815, App
28	453	16.0	1383	4	US-09-489-039A-2492	Sequence 2492, Ap
29	445.5	15.7	1074	4	US-09-270-767-13957	Sequence 13957, A
30	420	14.8	3486	4	US-09-976-594-1018	Sequence 1018, Ap
31	410	14.5	63783	4	US-09-949-016-13576	Sequence 13576, A
32	375	13.2	976	4	US-09-270-767-13814	Sequence 13814, A
33	374	13.2	217	4	US-09-016-434-602	Sequence 602, App
34	373	13.2	1305	4	US-09-489-039A-6462	Sequence 6462, Ap
35	371.5	13.1	9334	3	US-09-172-952-19	Sequence 19, Appl
36	370	13.0	1413	4	US-09-328-352-2786	Sequence 2786, Ap
37	369	13.0	1317	3	US-09-172-952-5	Sequence 5, Appli
38	369	13.0	1500	4	US-09-922-501-11	Sequence 11, Appl
39	368	13.0	471	4	US-09-270-767-30032	Sequence 30032, A
40	355.5	12.5	1368	4	US-09-489-039A-2771	Sequence 2771, Ap
41	346.5	12.2	1380	4	US-09-489-039A-3459	Sequence 3459, Ap
42	332	11.7	492	4	US-09-270-767-1297	Sequence 1297, Ap
43	332	11.7	492	4	US-09-270-767-16579	Sequence 16579, A
44	318.5	11.2	1269	4	US-09-891-641-65	Sequence 65, Appl
45	313.5	11.1	1332	4	US-09-489-039A-2164	Sequence 2164, Ap

ALIGNMENTS

RESULT 1
US-09-359-167-7
; Sequence 7, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/01US
; CURRENT APPLICATION NUMBER: US/09/359,167
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: 60-693,843
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (263)..(1870)
US-09-359-167-7

Alignment Scores:
Pred. No.: 3e-300 Length: 2930
Score: 2836.00 Matches: 536
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-776-865-2 (1-536) x US-09-359-167-7 (1-2930)

Qy 1 MetAlaAlaClyAlaMetThrProProArgProValGlnProAlaArgProGlyGlyPhe 20

Db 263 ATGGGGGGGGGGGATGACACCGCCCGCGGTCCAGCCAGCTCGCGCGGGGGCTTC 322
Qy 21 GlyLeuSerGlyArgSerLeuLeuCysGlnValAlaSerThrProAlaHisValGly 40
Db 323 GGGCTGTGGCGGGCGCTCCCTCTCTGCCAGGTGGGAGTACACCTGCTCAGTAGGC 382
Qy 41 ValMetArgSerProValArgAspLeuAlaArgAsnAspGlyGluGluSerThrAspArg 60
Db 383 GTCATGAGGTCTCGGGTTCGAGACCTGGCCCGGAACGATGGCAGGAGAGCAGCGACCGC 442
Qy 61 ThrProLeuLeuProGlyAlaProArgAlaGluAlaProValCysCysSerAlaArg 80
Db 443 ACGCTCTTTCACCGGGCGCCCGCCAGCGCCGCTCAGGTGTCTGCTCTGCTCGT 502
Qy 81 TyrAsnLeuAlaLeuAlaPhePheGlyPhePheLeuValTyrAlaLeuArgValAsn 100
Db 503 TACAACCTAGCAATTTGGCCCTTTTGGTTCTTCATTGTGTATGCATTACGTGTGAAT 562
Qy 101 LeuSerValAlaLeuValAspMetValAspSerAsnThrThrLeuGluAspAsnArgThr 120
Db 563 CTGAGTGTGGGTAGTGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 622
Qy 121 SerIysAlaCysProGluHisSerAlaProIleLysValHisAsnGlnThrGlyLys 140
Db 623 TCCAAGGGCGTTCAGAGCATTCCTCCCAATAAAGTTTCATCAATCAAAACGGGTAA 682
Qy 141 LysTyrGlnTrpAspAlaGluThrGlnGlyTrpIleLeuGlySerPhePheTyrGlyTyr 160
Db 683 AAGTACCANTGGGATGCGAAGCTCAAGATGGATTCCTGGTTCTCTTTTATGGCTAC 742
Qy 161 IleIleThrGlnIleProGlyGlyTyrValAlaSerLysIleGlyLysMetLeuLeu 180
Db 743 ATCATCACACAGATTCCTGGAGGATATGTTGCCAGCAAAATAGGGGGGAAAATGCTGTA 802
Qy 181 GlyPheGlyIleLeuGlyThrAlaValLeuThrLeuPheThrProIleAlaAlaAspLeu 200
Db 803 GGATTTGGGATTCCTGGGACATGCTGCTCCACCTGTTCACTCCCATTCCTGCGAGATTA 862
Qy 201 GlyValGlyProLeuIleValLeuArgAlaLeuGluGlyLeuGlyGluValThrPhe 220
Db 863 CGAGTTGGACCACTCATTTGTTACTCAGACCATAGAGCACTAGAGAGGGTGTACATTT 922
Qy 221 ProAlaMetHisAlaMetTrpSerSerTrpAlaProProLeuGluArgSerLysLeuLeu 240
Db 923 CCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 982
Qy 241 SerIleSerTyrAlaGlyAlaGlnLeuGlyThrValIleSerLeuProLeuSerGlyIle 260
Db 983 AGCATTTTCGTATGCAGGAGCAGCTGGGACAGTAATTTCTCTCTCTTTCTGGAATA 1042
Qy 261 IleCysTyrTyrMetAsnTrpThrTyrValPheTyrPheGlyThrIleGlyIlePhe 280
Db 1043 ATTGTCTACTATGATGAATGACATTAATGCTCTCTCTCTCTCTCTCTCTCTCTCT 1102
Qy 281 TrpPheLeuLeuTrpIleTrpLeuValSerAspThrProGlnLysHisLysArgIleSer 300
Db 1103 TGGTTTCTTTTGTGATCTGGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 1162
Qy 301 HisTyrGluLysGluTyrIleLeuSerSerLeuArgAsnGlnLeuSerSerGlnLysSer 320
Db 1163 CATTTATGAAAGGAATAATCTTTTCATCATTAAGAAATCAGCTTTCTTTCACAGAGTCA 1222
Qy 321 ValProTrpValProIleLeuLysSerLeuProLeuTrpAlaIleValValAlaHisPhe 340
Db 1223 GTGCGGTGGGTACCAATTTTAAATCCCTGTCACCTTTGGGCTATGCTAGTTGCACTTT 1282
Qy 341 SerTyrAsnTrpThrPheTyrThrLeuLeuThrLeuLeuProThrTyrMetLysGluIle 360
Db 1283 TCTTACAACCTGGACTTTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCT 1342
Qy 361 LeuArgPheAsnValGlnGluAsnGlyPheLeuSerSerLeuProTyrLeuGlySerTrp 380
Db 1343 CTAAGGTTCAATGTTCAGAGATGGGGTTTTTATCTCTCATTTGCCTTATTTAGGCTCTTGG 1402

Qy 381 LeuCysMetIleLeuSerGlyGlnAlaAlaAspAsnLeuArgAlaLysTrpAsnPheSer 400
Db 1403 TTATGTATGATCTCTGTCTCAAGCTGCTGCAAAATTAAGGGCAAAATGGAATTTTCA 1462
Qy 401 ThrLeuCysValArgArgIlePheSerLeuIleGlyMetIleGlyProAlaValPheLeu 420
Db 1463 ACTTTATGTCTTCGCAAAATTTTATAGCTTTATAGGAATGATGGACCTGAGTATTCCTG 1522
Qy 421 ValAlaAlaGlyPheIleGlyCysAspTyrSerLeuAlaValAlaPheLeuThrIleSer 440
Db 1523 GTAGCTGCTGGCTTCATGCTGTGATTAATCTTTGGCCGCTTCTTCTTAACTATATCA 1582
Qy 441 ThrThrLeuGlyGlyPheCysSerSerGlyPheSerIleAsnHisLeuAspIleAlaPro 460
Db 1583 ACAACCTGGAGGCTTTTCTCTCTGATTTAGCATCAACCATCTGGATATTTGCTCTCT 1642
Qy 461 SerTyrAlaGlyIleLeuLeuGlyIleThrAsnThrPheAlaThrIleProGlyMetVal 480
Db 1643 TCGTATGCTGCTATCTCTCTGGCATCAAAATACATTTTCCCATTAATTCAGGAATGTT 1702
Qy 481 GlyProValIleAlaLysSerLeuThrProAspAsnThrValGlyGluTrpGlnThrVal 500
Db 1703 GGGCCCGCTCATTTGCTAAAGTCTGACCCCTGATAACACTGTTGGAGNATGGCAACCTG 1762
Qy 501 PheTyrIleAlaAlaAlaIleAsnValPheGlyAlaIlePhePheThrLeuPheAlaLys 520
Db 1763 TTTCTATATTGCTGCTATTAATGTTTTTGGTGCCATTTTCTTACACTATTTCGCCAAA 1822
Qy 521 GlyGluValGlnAsnTrpAlaLeuAsnAspHisHisGlyHisArgHis 536
Db 1823 GGTGAAGTACAAAACCTGGGCTCTCAATGATCACCATGGACACAGACAC 1870

RESULT 2
US-09-949-016-1834
; Sequence 1834, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1834
; LENGTH: 2513
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1834

Alignment Scores:
Pred. No.: 6,62e-297 Length: 2513
Score: 2804.50 Matches: 534
Percent Similarity: 96.05% Conservative: 1
Best Local Similarity: 95.87% Mismatches: 1
Query Match: 98.89% Indels: 21
DB: 4 Gaps: 1

US-09-776-865-2 (1-536) x US-09-949-016-1834 (1-2513)
Qy 1 MetAlaAlaGlyAlaMetThr----- 7
Db 86 ATGGCGGGGGGGGGAGTCTGGCGCGCGCTTCTGGCGGGGACCGGGGGACTAGAGT 145
Qy 8 -----ProArgProValGlnProAlaArgProGlyGly 19

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2005, 02:43:23 ; Search time 803.74 Seconds
(without alignments)
3947.771 Million cell updates/sec

Title: US-09-776-865-2

Perfect score: 2836

Sequence: 1 MAAGAMTPRPVQPARPGF.....LFAKEGVQNALNDHGHGRH 536

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO_spool/US09776865/runat 08072005 175548 23365/app query.fasta 1.1358
-DB=N Geneseq 16Dec04 -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09776865 @CGN 1 1 1004 @runat 08072005 175548 23365 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04.*

1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2836	100.0	2930	3 AAZ50879	Aaz50879 Full leng
2	2836	100.0	2930	4 AAD10325	Aad10325 Human gro
3	2804.5	98.9	2511	12 ADQ84158	Adq84158 Human tum
4	2804	98.9	2512	4 AAF55900	Aaf55900 Human AST
5	2804	98.9	2626	11 ACN91332	Acn91332 Breast ca

6	2804	98.9	3362	5 ADL45207	Adl45207 Human ova
7	2770.5	97.7	3329	12 ADJ75057	Adj75057 Marker ge
8	2770.5	97.7	3329	13 ADRI4586	Adri4586 Human NF-
9	2770.5	97.7	3329	13 ADP25216	Adp25216 PRO poly
10	2734	96.4	2712	4 AAK94876	Aak94876 Human ful
11	2734	96.4	2712	12 ADL32035	Adl32035 Full leng
12	2717	95.8	2602	3 AAZ50875	Aaz50875 Partial h
13	2614	92.2	1488	4 AAI58115	Aai58115 Human pol
14	2338	82.4	2844	3 AAZ50876	Aaz50876 Sheep GBS
15	2338	82.4	2844	4 AAD10326	Aad10326 Sheep gro
16	2217	78.2	1485	3 AAZ50880	Aaz50880 Human/She
17	1967	69.4	1485	3 AAZ50881	Aaz50881 Human/She
18	1716	60.5	2670	4 AAH79234	Aah79234 Human sod
19	1614	56.9	929	4 AAI59901	Aai59901 Human pol
20	1521	53.6	1872	6 ABQ54422	Abq54422 Human ova
21	1515	53.4	1651	6 ABL90384	Abi90384 Human pol
22	1417	50.0	853	4 AAK93901	Aak93901 Human con
23	1417	50.0	853	12 ADL30328	Adl30328 3' end of
24	1339.5	47.2	838	4 AAK92364	Aak92364 Human cDN
25	1339.5	47.2	838	12 ADL48791	Adl28791 5' end of
26	996.5	35.1	1939	4 ABL03769	Abi03769 Drosophil
27	992	35.0	1641	4 ABL18113	Abi18113 Drosophil
28	992	35.0	1786	4 ABL07417	Abi07417 Drosophil
29	984.5	34.7	4899	4 ABL03768	Abi03768 Drosophil
30	973	34.3	3982	8 ABX13554	Abx13554 Rat DNPI
31	973	34.3	3982	9 ACF25330	Acf25330 Rat Na-de
32	973	34.3	3982	10 ADC15491	Adc15491 Rat DNPI
33	971	34.2	2528	8 ABX13555	Abx13555 Murine DN
34	971	34.2	2528	10 ADC15493	Adc15493 Mouse DNP
35	968	34.1	3946	8 ABX13553	Abx13553 Human DNP
36	968	34.1	3946	10 ADC15489	Adc15489 Human DNP
37	962	33.9	3422	4 AAK52406	Aak52406 Human pol
38	956	33.7	1811	6 ABK87049	Abk87049 Human tra
39	956	33.7	1811	10 ADG88328	Adg88328 Human tra
40	955	33.7	1767	10 ADD01393	Add01393 Human TCH
41	955	33.7	1850	10 ADD01409	Add01409 Human TCH
42	955	33.7	3671	9 AAL62505	Aal62505 Human tra
43	955	33.7	3952	13 ADR08065	Adr08065 Full leng
44	945	33.3	1803	10 ADD01411	Add01411 Mouse TCH
45	945	33.3	1822	10 ADD01422	Add01422 Mouse TCH

ALIGNMENTS

RESULT 1

AAZ50879

ID AAZ50879 standard; cDNA; 2930 BP.

XX AAZ50879;

XX 31-MAY-2000 (first entry)

Full length human GBS toxin receptor (HP59) cDNA.

Human GBS toxin receptor; group B beta-haemolytic streptococci; HP59; pathological vascularisation; cancer metastases; angiogenesis; neovascularisation; reperfusion injury; scarring; keloid; chronic inflammation; rheumatoid arthritis; psoriasis; neural injury; endothelial cell proliferation; antibacterial; anticancer; anti-angiogenic; anti-inflammatory; anti-arthritis; anti-psoriatic; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 263..1873

XX FT /**tag= a

XX FT /product= "Human GBS toxin receptor protein"

XX WO200005375-A1.

XX PD 03-FEB-2000.

XX PF 22-JUL-1999; 99WO-US016676.

```
XX 22-JUL-1998; 98US-0093843P.
XX (UYVA-) UNIV VANDERBILT.
XX Heilerqvist CG, Fu C;
XX WPI; 2000-205377/18.
XX P-PSDB; AAY45089.
XX New polynucleotide encoding mammalian receptor for streptococcus toxin,
XX useful for diagnosis and treatment of, e.g. pneumonia in neonates.
XX Claim 3; Page 89-93; 109pp; English.
XX The present cDNA sequence encodes full length human GBS (Group B beta -
XX haemolytic streptococci) toxin receptor (HP59). This sequence was cloned
XX by using human embryo lung cDNA library as template. Expression vectors
XX comprising this cDNA can be transformed into host cells to express GBS
XX toxin receptor and its fragments. Detecting the receptor in tissues is
XX used to diagnose pathological vascularisation, e.g. for detecting cancer
XX metastases. GBS toxin receptors are useful for treating conditions
XX associated with pathological angiogenesis or neovascularisation
XX (specifically cancer, reperfusion injury, scarring during wound healing,
XX keloids, chronic inflammation (rheumatoid arthritis or psoriasis) or
XX neural injury), and to raise specific antibodies used for treating early
XX onset disease. Inhibitors of this receptor are useful for treating
XX pathological or hypoxia-induced endothelial cell proliferation and
XX migration
XX
XX SQ Sequence 2930 BP; 729 A; 698 C; 681 G; 822 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 4.4e-264 Length: 2930
XX Score: 2836.00 Matches: 536
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 3 Gaps: 0
XX
XX US-09-776-865-2 (1-536) x AAZ50879 (1-2930)
XX
XX 1 MetAlaAlaGlyAlaMetThrProArgProValGlnProAlaArgProGlyGlyPhe 20
XX
XX 263 ATGGCGGGGGGGGATGACACCGCCCGCGGTCCAGCCAGCTCGCGCGGGGGCTTC 322
XX
XX 21 GlyLeuSerGlyArgSerLeuLeuCysGlnValAlaSerThrProAlaHisValGly 40
XX
XX 323 GGGCTGTCTGGGGCGGCGCTCCCTTCTCTGCCAGGTGGCGAGTACACCTGCTCAGGTAGGC 382
XX
XX 41 ValMetArgSerProValArgAspLeuAlaArgAsnAspGlyGluGluSerThrAspArg 60
XX
XX 383 GTCATGAGTCTCCGGTTCGAGACTGGCCCGGAGACGATGGCGAGGAGACGCGACCGC 442
XX
XX 61 ThrProLeuLeuProGlyAlaProArgAlaGluAlaProValCysCysSerAlaArg 80
XX
XX 443 ACGGCTCTTCTACGGGGCGGCCACCGGGCCGAGCGCTCCAGGTGTGCTGCTCGT 502
XX
XX 81 TyrAsnLeuAlaLeuLeuAlaPhePheGlyPhePheLeuValTyrAlaLeuArgValAsn 100
XX
XX 503 TACAACCTTAGCAATTTTGGCCCTTTTGGTTTCTTCATTGTGTATGATACGTGTGAAT 562
XX
XX 101 LeuSerValAlaLeuValAspMetValAspSerAsnThrThrLeuGluAspAsnArgThr 120
XX
XX 563 CTGAGTGTGGGTAGTGATATGATGATGATGATGATGATGATGATGATGATGATGAT 622
XX
XX 121 SerLysAlaCysProGluHisSerAlaProLeuLysValHisHisAsnGlnThrGlyLys 140
XX
XX 623 TCCAAGGGGTCTCCAGAGCATTCCTGCCATAAAGATTCATCATATCAATCAACCGGTAAG 682
XX
XX 141 LysTyrGlnTrpAspAlaGluThrGlnGlyTrpLeuLeuGlySerPhePheTyrGlyTyr 160
XX
XX 683 AAGTACCAATGGGATGCGAGAACTCAAGGATGGATTCCTCGGTTCCTTTTATGGGTAC 742
```

```
Qy 161 IleIleThrGlnIleProGlyGlyTyrValAlaSerLysIleGlyGlyLysMetLeuLeu 180
Db 743 ATCATCACAGATTCCTGGAGGATATGTTGCCAGCAAAATAGGGGGGAAAATGCTCTCTA 802
Qy 181 GlyPheGlyIleLeuGlyThrAlaValLeuThrLeuPheThrProIleAlaAlaAspLeu 200
Db 803 GGATTTGGGATCCTTGGCACTGCTGTCTCCTCACCTGTTCACCTCCCATTCGTGTCAGATTTA 862
Qy 201 GlyValGlyProLeuIleValLeuArgAlaLeuGluGlyLeuGlyGluGlyValThrPhe 220
Db 863 GGAGTTGGACCATCATTTGTTACTCAGAGCACTAGAGGACTAGAGAGGGGTGTACATTT 922
Qy 221 ProAlaMetHisAlaMetTrpSerSerTrpAlaProProLeuGluArgSerLysLeuLeu 240
Db 923 CAGCCATGATGCATGCGATGTGCTTCTTGGGCTCCCTCTTGAAGAAGCAAACTTCTT 982
Qy 241 SerIleSerTyrAlaGlyAlaGlnLeuGlyThrValIleSerLeuProLeuSerGlyIle 260
Db 983 AGCATTTGCTATCGAGGAGCACAGCTTGGGACAGTAAATTTCTCTCTCTTTCTGGAATA 1042
Qy 261 IleCysTyrTyrMetAsnTrpThrTyrValPheTyrPhePheGlyThrIleGlyIlePhe 280
Db 1043 ATTTGCTACTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1102
Qy 281 TrpPheLeuLeuTrpIleTrpLeuValSerAspThrProGlnLysHisLysArgIleSer 300
Db 1103 TGGTTTCTTTTGGGATCTGGTTAGTGTAGTGCACACCAACCAACCAAGAGAAATTTCC 1162
Qy 301 HisTyrGlnLysGluTyrIleLeuSerSerLeuArgAsnGlnLeuSerSerGlnLysSer 320
Db 1163 CATTTATGAAAGGAATACATTTCTTATCAATTAAGAAATCAGCTTTCTTCCACAGAACTCA 1222
Qy 321 ValProTrpValProIleLeuLysSerLeuProLeuTrpAlaIleValValAlaHisPhe 340
Db 1223 GTGCCGTGGGTACCCATTTTAAATCCCTGCCACTTGGGCTATCGTAGTTGCACACTTT 1282
Qy 341 SerTyrAsnTrpThrPheTyrThrLeuLeuThrLeuLeuProThrTyrMetLysGluIle 360
Db 1283 TCTTACAACTGGAGCTTTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 1342
Qy 361 LeuArgPheAsnValGlnGluAsnGlyPheLeuSerSerLeuProTyrLeuGlySerTrp 380
Db 1343 CTAAGGTTCATGTTCAAGAGAAATGGGTTTATCTTATCTTATCTTATCTTATCTTATCT 1402
Qy 381 LeuCysMetIleLeuSerGlyGlnAlaAlaAspAsnLeuArgAlaLysTrpAsnPheSer 400
Db 1403 TTAATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462
Qy 401 ThrLeuCysValArgArgIlePheSerLeuIleGlyMetIleGlyProAlaValPheLeu 420
Db 1463 ACTTTATGTTTCGAGAAATTTTATGCTTATAGGAATGATGAGCTGCTGCTGCTGCTGCT 1522
Qy 421 ValAlaAlaGlyPheIleGlyCysAspTyrSerLeuAlaValAlaPheLeuThrIleSer 440
Db 1523 GTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1582
Qy 441 ThrThrLeuGlyGlyPheCysSerSerGlyPheSerIleAsnHisLeuAspIleAlaPro 460
Db 1583 ACAACACTGGAGGCTTTTGTCTCTTCTGGATTTAGCATCAACCATCTGGATATGCTCTCT 1642
Qy 461 SerTyrAlaGlyIleLeuLeuGlyIleThrAsnThrPheAlaThrIleProGlyMetVal 480
Db 1643 TCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1702
Qy 481 GlyProValIleAlaLysSerLeuThrProAspAsnThrValGlyGluTrpGlnThrVal 500
Db 1703 GGCCCCGCTCATGCTAAAGCTGACCCCTGATAACACTGTTGGAGAAATGGCAACCGCTG 1762
Qy 501 PheTyrIleAlaAlaAlaIleAsnValPheGlyAlaIlePhePheThrLeuPheAlaLys 520
Db 1763 TTCTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1822
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 9, 2005, 02:51:19 ; Search time 6822.95 seconds
(without alignments)
3806.563 Million cell updates/sec

Title: US-09-776-865-2
Perfect score: 2836
Sequence: 1 MAAGMTPPRVQPARPGF.....LPAXGEVQNWALNDHGHHRH 536

Scoring table: BIOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09776865/runat.08072005.175549.23371/app.query.fasta.1.1358
-DB=GenEmbl -QFMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=spt -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09776865 @CGN 1.1 8225 @runat.08072005.175549.23371 -NCPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.on.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.scs.*
12: gb.by.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2836	100.0	2930	6	BD248129
2	2836	100.0	2930	6	AX207624
3	2836	100.0	2930	9	AF244577
4	2804	98.9	2512	6	AX138494

5	2804	98.9	2512	9	HSA387747	9	AD387747	Homo sapi
6	2804	98.9	3362	6	C0412026	6	C0412026	Sequence
7	2770.5	97.7	3329	6	CQ77623	6	CQ77623	Sequence
8	2761	97.4	3292	9	BC020961	9	BC020961	Homo sapi
9	2734	96.4	2712	6	CQ783928	6	CQ783928	Sequence
10	2734	96.4	2712	6	BD127905	6	BD127905	Primer fo
11	2734	96.4	2712	9	AK075320	9	AK075320	Homo sapi
12	2717	95.8	2602	6	BD248125	6	BD248125	GBS toxin
13	2338	82.4	2844	4	AF244578	4	AF244578	Ovis arie
14	2338	82.4	2844	6	BD248126	6	BD248126	GBS toxin
15	2338	82.4	2844	6	AX207626	6	AX207626	Sequence
16	2217	78.2	1485	6	BD248130	6	BD248130	GBS toxin
17	2145	75.6	3152	10	BC058785	10	BC058785	Mus muscu
18	2078.5	73.3	3121	5	AJ719840	5	AJ719840	Gallus ga
19	1967	69.4	1485	6	BD248131	6	BD248131	GBS toxin
20	1417	50.0	853	6	CQ782221	6	CQ782221	Sequence
21	1417	50.0	853	6	BD126930	6	BD126930	Primer fo
22	1339.5	47.2	838	6	CQ780684	6	CQ780684	Sequence
23	1339.5	47.2	838	6	BD125393	6	BD125393	Primer fo
24	1183	41.7	1963	3	AK114957	3	AK114957	Ciona int
25	1143.5	40.3	2019	3	AK114801	3	AK114801	Ciona int
26	996.5	35.1	1939	6	CQ575136	6	CQ575136	Sequence
27	992	35.0	1641	6	CQ596652	6	CQ596652	Sequence
28	992	35.0	1786	6	CQ580608	6	CQ580608	Sequence
29	992	35.0	1841	3	AY060776	3	AY060776	Drosophil
30	986	34.8	2085	3	BT010092	3	BT010092	Drosophil
31	984.5	34.7	4899	6	CQ575135	6	CQ575135	Sequence
32	984.5	34.7	3344	2	AC014246	2	AC014246	Drosophil
33	984.5	34.7	167926	3	AC023685	3	AC023685	Drosophil
34	984.5	34.7	174157	3	AC023711	3	AC023711	Drosophil
35	984.5	34.7	332029	3	AE003491	3	AE003491	Drosophil
36	973	34.3	3982	6	AX700122	6	AX700122	Sequence
37	973	34.3	3982	6	AX709534	6	AX709534	Sequence
38	973	34.3	3982	6	AX743494	6	AX743494	Sequence
39	973	34.3	3982	10	AF271235	10	AF271235	Rattus no
40	971	34.2	2528	6	AX709538	6	AX709538	Sequence
41	971	34.2	2528	6	AX743498	6	AX743498	Sequence
42	971	34.2	2528	10	AF324864	10	AF324864	Mus muscu
43	971	34.2	3728	10	BC038375	10	BC038375	Mus muscu
44	968	34.1	2019	9	BC069629	9	BC069629	Homo sapi
45	968	34.1	2020	9	BC069640	9	BC069640	Homo sapi

ALIGNMENTS

RESULT 1	BD248129	BD248129	GBS toxin receptor.	2930 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD248129	BD248129	GBS toxin receptor.				
DEFINITION	BD248129	BD248129	GBS toxin receptor.				
ACCESSION	BD248129	BD248129	GBS toxin receptor.				
VERSION	BD248129.1	GI:33057899					
KEYWORDS	JP 2002524027-A/5.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
REFERENCE	1	(bases 1 to 2930)					
AUTHORS	Hellerqvist, C.G. and Fu, C.						
TITLE	GBS toxin receptor						
JOURNAL	VANDERBILT UNIVERSITY						
COMMENT	OS Homo sapiens (human)						
	PN JP 2002524027-A/5						
	PD 06-AUG-2002						
	PR 22-JUL-1999	JP 2000561321					
	PI CARL G HELLERQVIST, CHANGLIN FU	60/093843					
	PC C12N15/09, A61K45/00, A61P9/00, A61P17/02, A61P25/00, A61P29/00, PC						
	A61P35/00,						
	PC A61P43/00, C07K14/705, C07K16/28, C12N1/15, C12N1/19, C12N1/21, PC						
	C12N5/10,						
	PC C12P21/02, C12Q1/02, G01N33/15, G01N33/50, G01N33/50, C12N5/00 CC						
	GBS toxin receptor						

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 05:13:45 ; Search time 106.106 Seconds

(without alignments)
2388.931 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 2617

Sequence: 1 MKSPVSLAPSDGEGSDRT.....LFKAGEVQNVAISHQGRN 495

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2617	100.0	495	1 S175 SHEEP	Q9mzdl ovis aries
2	2329	89.0	495	1 S175 HUMAN	Q9nra2 homo sapien
3	2236	85.4	495	1 S175 MOUSE	Q8bn82 mus musculu
4	1058	40.4	544	2 Q7Q580	Q7q580 anopheles g
5	1006	38.4	476	2 Q7Q579	Q7q579 anopheles g
6	991.5	37.9	559	2 Q9VYG7	Q9vyg7 drosophila
7	985.5	37.7	502	2 Q3VDM0	Q9vdm0 drosophila
8	979	37.4	589	2 Q8NDX2	Q8ndx2 homo sapien
9	976	37.3	588	2 Q8KIQ1	Q8kiq1 rattus norv
10	976	37.3	588	2 Q7TSF2	Q7tsf2 rattus norv
11	969.5	37.0	601	2 Q8BFU8	Q8bfu8 mus musculu
12	968.5	37.0	493	1 YLD2 CAEL	Q3567 caenorhabdi
13	965	36.9	582	2 Q920E7	Q920b7 mus musculu
14	965	36.9	582	2 Q8TII2	Q8jil2 rattus norv
15	962.5	36.8	582	2 Q9P2U8	Q9p2u8 homo sapien
16	960	36.7	582	2 Q8BLE7	Q8ble7 mus musculu
17	941	36.0	529	1 P1CO DROME	Q9v785 drosophila
18	940.5	35.9	576	2 Q6INC8	Q6inc8 xenopus lae
19	940	35.9	584	2 Q8AW47	Q8aw47 brachydanio
20	936	35.8	479	2 Q23514	Q23514 caenorhabdi
21	932	35.6	560	2 Q9P2U7	Q9p2u7 homo sapien
22	927	35.4	560	2 Q62634	Q62634 rattus norv
23	926	35.4	483	1 P1CO DROAN	Q61369 drosophila
24	923	35.3	560	2 Q6PCD0	Q6pcd0 homo sapien
25	899.5	34.4	576	2 Q715L3	Q715l3 xenopus lae
26	884	33.8	466	2 Q96LH1	Q96lh1 homo sapien
27	881	33.7	955	2 Q932C5	Q9367 anopheles g
28	874.5	33.4	497	2 Q7Y2C5	Q7y2c5 homo sapien
29	860	32.9	535	2 Q7PWK4	Q7pwk4 anopheles g
30	859.5	32.8	481	2 Q9VJW8	Q9vjw8 drosophila
31	849.5	32.5	533	2 Q7Q1S5	Q7q1s5 anopheles g

32	847.5	32.4	453	2	Q9NKF8	Q9nkf8 drosophila
33	845	32.3	479	2	Q9W4G7	Q9w4g7 drosophila
34	841.5	32.2	498	2	Q8CJH9	Q8cjh9 rattus norv
35	835.5	31.9	496	2	Q9VKC9	Q9vkc9 drosophila
36	833.5	31.8	470	2	Q7YTZ7	Q7ytz7 drosophila
37	819	31.3	563	2	Q9TZN7	Q9tzn7 caenorhabdi
38	819	31.3	576	1	YQO6 CAEL	P34644 caenorhabdi
39	812.5	31.0	485	2	Q7PSC6	Q7psc6 anopheles g
40	809.5	30.9	465	1	NPT1 RABIT	Q28722 o renal sod
41	807.5	30.9	465	1	Q9IWW5	Q9iww5 mus musculu
42	806.5	30.8	498	2	Q8VCY5	Q8vcy5 mus musculu
43	804.5	30.7	467	2	Q7YQJ7	Q7yqj7 sus scrofa
44	803	30.7	447	2	Q8VCX8	Q8vcx8 mus musculu
45	802.5	30.7	465	1	NPT1 MOUSE	Q61983 mus musculu

ALIGNMENTS

RESULT 1

S175_SHEEP STANDARD; PRT; 495 AA.
ID S175_SHEEP
AC Q9WZD1;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Sialin (Solute carrier family 17 member 5) (Sodium/sialic acid cotransporter) (Membrane glycoprotein SP55).
GN Name=SLC17A5;
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE=Lung;
RX PubMed=11751519;
RA Fu C., Hardhan S., Cetateanu N.D., Wamil B.D., Wang Y., Yan H.-P., Shi E., Carter C., Venkov C., Yakes F.M., Page D.L., Lloyd R.S., Mernaugh R.L., Hellerqvist C.G.;
RT "Identification of a novel membrane protein, HP59, with therapeutic potential as a target of tumor angiogenesis.";
RL Clin. Cancer Res. 7:4182-4194(2001).
CC FUNCTION: Primary solute translocator for anionic substances; particularly it is a free sialic acid transporter in the lysosomes (Probable). Receptor for CM101, a polysaccharide produced by group B Streptococcus with antipathogenic properties.
CC B Streptococcus with antipathogenic properties.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Lysosomal (Potential).
CC -!- TISSUE SPECIFICITY: Significantly expressed in lung endothelial cells, and much less in liver.
CC -!- SIMILARITY: Belongs to the major facilitator (MFS) superfamily. Sodium/anion cotransporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF244578; AAF97770.1; -;
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Subtransporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PSS0850; MFS; 1.
DR Lysosome; Receptor; Symport; Transmembrane; Transport.
KW TRANSMEM 42 62 Potential.
FT TRANSMEM 110 130 Potential.
FT TRANSMEM 137 157 Potential.
FT TRANSMEM 159 179 Potential.

FT TRANSMEM 201 221 Potential.
FT TRANSMEM 228 248 Potential.
FT TRANSMEM 289 309 Potential.
FT TRANSMEM 329 349 Potential.
FT TRANSMEM 366 386 Potential.
FT TRANSMEM 392 412 Potential.
FT TRANSMEM 424 444 Potential.
FT TRANSMEM 458 478 Potential.
SQ SEQUENCE 495 AA; 54536 MW; 649D7C4A59B28272 CRC64;

Query Match 100.0%; Score 2617; DB 1; Length 495;
Best Local Similarity 100.0%; Pred. No. 2.8e-183;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSPVSLAPSDGEGSDRTLLQAPRAEPAPVCCSARYNLAFSPFFVLSLRLVNL 60
Db 1 MKSPVSLAPSDGEGSDRTLLQAPRAEPAPVCCSARYNLAFSPFFVLSLRLVNL 60

Qy 61 SVALVDMVDSNTAKDNRTSYECAHSAPIKVLHNQTKGYRDAETQGWILGSFFYGYI 120
Db 61 SVALVDMVDSNTAKDNRTSYECAHSAPIKVLHNQTKGYRDAETQGWILGSFFYGYI 120

Qy 121 ITQIPGGVVASRSGKLLGFGIFATAFTFTPLAADFGVGVALVRLALEGLGEGVTYP 180
Db 121 ITQIPGGVVASRSGKLLGFGIFATAFTFTPLAADFGVGVALVRLALEGLGEGVTYP 180

Qy 181 AHWAMSSWAPLERSKLLSISYAGAQLGTVVSLPSGVICYNNWTYVFFFGIVGLIIV 240
Db 181 AHWAMSSWAPLERSKLLSISYAGAQLGTVVSLPSGVICYNNWTYVFFFGIVGLIIV 240

Qy 241 FILWICLVSDTPETHKTTTPYEKEYILSSLNQNLSSQKSPWIPMLKSLPLWAIWVAFHS 300
Db 241 FILWICLVSDTPETHKTTTPYEKEYILSSLNQNLSSQKSPWIPMLKSLPLWAIWVAFHS 300

Qy 301 YNWTFTYLLTTPYMKVLRNFQENGFLSAVPLGVCWLCWILSGQADNLRWNPST 360
Db 301 YNWTFTYLLTTPYMKVLRNFQENGFLSAVPLGVCWLCWILSGQADNLRWNPST 360

Qy 361 LWVRVFLSLIGMIGPAIFLVAAGFTGCDYSLAVAFITSTTLTGFCSSGFSINHLDIAPS 420
Db 361 LWVRVFLSLIGMIGPAIFLVAAGFTGCDYSLAVAFITSTTLTGFCSSGFSINHLDIAPS 420

Qy 421 YAGILLGTTNTFATIPGMIGPIIARSLTPENTIGEMQVTFCAAAINVFGAIFFTLPAKG 480
Db 421 YAGILLGTTNTFATIPGMIGPIIARSLTPENTIGEMQVTFCAAAINVFGAIFFTLPAKG 480

Qy 481 EVQNWAISSDHQHRN 495
Db 481 EVQNWAISSDHQHRN 495

RESULT 2

SI175_HUMAN STANDARD; PRT; 495 AA.
AC Q9NRA2; Q9NRS; Q9UGH0;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Sialin (Solute carrier family 17 member 5) (Sodium/sialic acid cotransporter) (AST) (Membrane glycoprotein HP59).
GN Name=SLC17A5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, AND TISSUE SPECIFICITY.
RX PubMed=11751519;
RA Fu C., Bardhan S., Cetateanu N.D., Wamil B.D., Wang Y., Van H.-P., Shi E., Carter C., Venkov C., Yakes F.M., Page D.L., Lloyd R.S., Mernaugh R.L., Hellerqvist C.G.;
RT "Identification of a novel membrane protein, HP59, with therapeutic potential as a target of tumor angiogenesis.";

RL Clin. Cancer Res. 7:4182-4194 (2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, TISSUE SPECIFICITY, VARIANT SD CYS-39, AND VARIANTS ISSD 268-SER--ASN-272 DEL; ARG-183 AND ARG-334.
RX MEDLINE=20047778; PubMed=10581036; DOI=10.1038/70585;
RA Verheijen F.W., Verbeek E., Aula N., Beerens C.E.M.T., Havelaar A.C., Joosse M., Peltonen L., Aula P., Galjaard H., Van der Spek P.J., Mancini G.M.S.;
RT "A new gene, encoding an anion transporter, is mutated in sialic acid storage diseases.";
RN Nat. Genet. 23:462-465 (1999).
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuna M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aoeuka S., Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Inagaki H., Ikema Y., Okamoto S., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Shigeta K., Senba T., Ohtani K., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human cDNAs.";
RN Nat. Genet. 36:40-45 (2004).
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Colon;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Trimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RP VARIANTS SD CYS-39 AND GLU-136, AND VARIANTS ISSD 268-SER--ASN-272 DEL; ARG-183; ARG-334 AND VAL-371.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:16:27 ; Search time 24.9661 Seconds
(without alignments)
1907.682 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 2617

Sequence: 1 MKSPVSLAPSDGEGSDRT.....LPKAGEVQNAISDHQGRN 495

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	968.5	37.0	493	2 G88553	protein C38C10.2 [
2	950	36.3	472	2 S28286	hypothetical prote
3	927	35.4	560	2 I59302	brain specific Na+
4	819	31.3	563	2 T43650	probable sodium-de
5	819	31.3	576	2 H88548	protein ZK512.6 [i
6	809.5	30.9	465	2 A56410	sodium/phosphate t
7	802.5	30.7	465	2 S59915	sodium-phosphate t
8	792.5	30.3	573	2 T23589	hypothetical prote
9	791.5	30.2	467	2 A48916	sodium phosphate t
10	775.5	29.6	465	2 I39473	Na+-dependent phos
11	766.5	29.3	544	2 T24633	hypothetical prote
12	744	28.4	512	2 H84698	hypothetical prote
13	585	22.4	501	2 B89135	protein F25G6.7 [i
14	561.5	21.5	413	2 T01534	hypothetical prote
15	550.5	21.0	537	2 T45634	hypothetical prote
16	541.5	20.7	466	2 S40767	hypothetical prote
17	499	19.1	561	2 D84800	hypothetical prote
18	490.5	18.7	592	2 T25419	hypothetical prote
19	489	18.7	568	2 S44742	CO2C2.4 protein -
20	471	18.0	530	2 T29418	hypothetical prote
21	450	17.2	380	2 T29248	hypothetical prote
22	444.5	17.0	516	2 T24729	hypothetical prote
23	428	16.4	499	2 T15201	hypothetical prote
24	420	16.0	506	2 T29688	hypothetical prote
25	419.5	16.0	543	2 T32496	hypothetical prote
26	395.5	15.1	493	2 T25357	hypothetical prote
27	391.5	15.0	445	2 T23590	hypothetical prote
28	382.5	14.6	425	2 A90055	hypothetical prote
29	379.5	14.5	473	2 T31717	hypothetical prote

30	373	14.3	455	1 H69752	probable glucarat
31	371	14.2	485	2 T24115	hypothetical prote
32	369	14.1	452	2 A80861	probable glucarate
33	368.5	14.1	478	2 T33942	hypothetical prote
34	366.5	14.0	462	2 T34365	hypothetical prote
35	365.5	14.0	420	2 S44900	ZK652.10 protein -
36	361	13.8	450	2 A65061	probable glucarate
37	358.5	13.7	428	2 AD0675	membrane transport
38	357	13.6	450	2 C85930	probable transport
39	357	13.6	450	2 A91085	probable transport
40	346.5	13.2	455	2 T34366	hypothetical prote
41	339.5	13.0	659	2 T33557	hypothetical prote
42	338.5	12.9	422	1 A69853	hexuronate transpo
43	334.5	12.8	493	2 T19383	hypothetical prote
44	331.5	12.7	516	2 T27092	hypothetical prote
45	326	12.5	487	2 T23776	hypothetical prote

ALIGNMENTS

RESULT 1

G88553

protein C38C10.2 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: G88553

R:anonymous, The C. elegans Sequencing Consortium.

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.eleg

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: G88553

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-493 <SFO>

A:Cross-references: UNIPROT:Q03567; GB:chr_III; PIDN:CAA79549.1; PID:G3874873; GSPDB:GN01

C:Genetics:

A:Gene: C38C10.2

A:Map position: 3

Query Match 37.0%; Score 968.5; DB 2; Length 493;

Best Local Similarity 40.5%; Pred. No. 1.9e-69;

Matches 193; Conservative 89; Mismatches 179; Indels 15; Gaps 6;

Qy	27	PRAPAPVCCSARYNLAFLSFFGFVLSRLVLSVALVDMVDSNTTAKONRTSYECAEH	86
Db	8	PRLPV-----STRLSLVNFVFGCLVTYMRTNMSFAVVCNVNENKTDGTGVEKYSRCGE	62
Qy	87	SAPIKVLHNTGKKYRWDATQGWILGSPFYGYITIQPGYVASRGGKLLGFGIPAT	146
Db	63	MTPVESNSSVIG-EFDMDKQTGMVLSSFFGYIGSQIIGHLASRYGGRVFWFTILGS	121
Qy	147	AIFTLFTPLAADFCVGLVALRALEGEGVTPAMHAMSSWAPPLERSKLLSISVAGA	206
Db	122	ALLTLNLPVAARTSEYALALIRAAIGFQATFPAMHTMMSVWGPPLELSVLTGVTYAGA	181
Qy	207	QLGTVWSLPSGVICYY---MNTVYVFYFGVIGIIFILWICLVSDTPTTHKTIITPYEK	263
Db	182	QIGNVIVLPSGLFCVCEYFGDGGWPSFIYIIGVFGVLTAVWWYSSDKPATHRITPEEK	241
Qy	264	EYILSSLNQLSSQ-----KSVPIPLMLKSLPLMAIVVAHFSYNWTFYLLTLTPTMKEV	319
Db	242	QYIVTAVEASMGKDTGKVPSTPIKILTSAPVACWAGHEAGDAGVATMLVSLPSFLKDV	301
Qy	320	LRFNIQENGFLSAVPLVGCWLCMLTSGAADNLBARWNFSTLWVRVFSGLIGMIPAIFL	379
Db	302	LGLNLSLGAVASIPYIAYFLAINAGVGLADTLRSKGLSTLNTTRRAAMLVALIGQGIFL	361
Qy	380	VAAGFICDYS-LAVAFLLTISTLLGGFCSSGFSINHLDIAPSYAGILLGINTFATIPGM	438
Db	362	VASGYCCGQDVLVIIITCGMAISGLQYAGVFNVYLEIAPFPFSGTVMGTNTISALAGI	421

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:22:07 ; Search time 106.586 seconds
(without alignments)
1794.174 Million cell updates/sec

Title: US-09-776-865-4
Perfect score: 2617
Sequence: 1 MKSPVSLAPSDGREGSDRT.....LFAKGEVQNAISDHQGRN 495

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1726216 seqs, 386330316 residues

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pap.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pap.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pap.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pap.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pap.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pap.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pap.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2617	100.0	495	9	US-09-776-865-4
2	2617	100.0	495	17	US-10-823-506-4
3	2329	89.0	495	9	US-09-915-181A-7
4	2329	89.0	495	16	US-10-755-889-588
5	2329	89.0	495	17	US-10-823-506-2
6	2329	89.0	495	9	US-09-776-865-2
7	2329	89.0	495	17	US-10-823-506-8
8	1702	65.0	495	17	US-10-823-506-10
9	1702	65.0	495	17	US-10-823-506-12
10	1386	53.0	284	15	US-10-264-049-2477
11	1318	50.4	272	15	US-10-264-237-2351

12	979	37.4	589	9	US-09-740-041-2	Sequence 2, Appli
13	979	37.4	589	14	US-10-389-967-2	Sequence 2, Appli
14	979	37.4	589	18	US-10-499-731-28	Sequence 28, Appli
15	969.5	37.0	601	18	US-10-499-731-46	Sequence 46, Appli
16	968.5	37.0	493	15	US-10-369-493-5834	Sequence 5834, Ap
17	965	36.9	578	9	US-09-740-041-4	Sequence 4, Appli
18	965	36.9	578	14	US-10-389-967-4	Sequence 4, Appli
19	965	36.9	582	9	US-09-915-181A-4	Sequence 4, Appli
20	965	36.9	582	15	US-10-205-331-7	Sequence 7, Appli
21	965	36.9	582	16	US-10-734-731-12	Sequence 12, Appli
22	965	36.9	582	16	US-10-734-731-14	Sequence 14, Appli
23	965	36.9	582	16	US-10-807-500-12	Sequence 12, Appli
24	965	36.9	582	16	US-10-807-500-14	Sequence 14, Appli
25	962.5	36.8	582	16	US-10-734-731-10	Sequence 10, Appli
26	962.5	36.8	582	16	US-10-807-500-10	Sequence 10, Appli
27	958	36.6	566	15	US-10-287-226-300	Sequence 300, App
28	956.5	36.5	850	9	US-09-915-181A-3	Sequence 3, Appli
29	932	35.6	560	16	US-10-734-731-2	Sequence 2, Appli
30	932	35.6	560	16	US-10-807-500-2	Sequence 2, Appli
31	931	35.6	560	15	US-10-314-790-5	GENERAL INFORMA
32	931	35.6	560	16	US-10-734-731-4	Sequence 4, Appli
33	931	35.6	560	16	US-10-807-500-4	Sequence 4, Appli
34	927	35.4	560	9	US-09-991-212A-4	Sequence 4, Appli
35	927	35.4	560	9	US-09-915-181A-5	Sequence 5, Appli
36	927	35.4	560	10	US-09-965-522-4	Sequence 4, Appli
37	927	35.4	560	16	US-10-734-731-6	Sequence 6, Appli
38	927	35.4	560	16	US-10-734-731-8	Sequence 8, Appli
39	927	35.4	560	16	US-10-807-500-6	Sequence 6, Appli
40	927	35.4	560	16	US-10-807-500-8	Sequence 8, Appli
41	927	35.4	560	17	US-10-877-818-4	Sequence 4, Appli
42	895	34.2	194	15	US-10-296-115-1200	Sequence 1200, Ap
43	819	31.3	563	9	US-09-915-181A-6	Sequence 6, Appli
44	819	31.3	576	14	US-10-014-079A-1	Sequence 1, Appli
45	819	31.3	576	15	US-10-369-493-5823	Sequence 5823, Ap

ALIGNMENTS

RESULT 1
US-09-776-865-4
; Sequence 4, Application US/09776865
; Patent No. US20020061846A1
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; TITLE OF INVENTION: Methods for Preventing or Attenuating Pathoangiogenic Conditions
; FILE REFERENCE: 22100-0100 46126-252687
; CURRENT APPLICATION NUMBER: US/09/776.865
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 60/179,870
; PRIOR FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Ovis sp.
US-09-776-865-4

Query Match	100.0%;	Score 2617;	DB 9;	Length 495;
Best Local Similarity	100.0%;	Pred. No. 9.3e-238;		
Matches 495;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MKSPVSLAPSDGEGSDRTPLQRA	PAPVCCSARYNLAFLSPFGFVLYSLRVNL	60
Db	1	MKSPVSLAPSDGEGSDRTPLQRA	PAPVCCSARYNLAFLSPFGFVLYSLRVNL	60
Qy	61	SVALVDVDSNTTAKONRTSYECAH	SAPIKVLHNQTGKKYRMDAETQGWILGSFFYGYI	120
Db	61	SVALVDVDSNTTAKONRTSYECAH	SAPIKVLHNQTGKKYRMDAETQGWILGSFFYGYI	120
Qy	121	ITQIPGGYVASRSGKLLGFGIFATA	IFTLFTPLAADFGVALVALRALEGEGGVTPP	180

Db 121 ITQIPGGVASSRGKLLGFGIFATAITFTPLAADFGVALVALRALEGLGEGVTYP 180
Qy 181 AMHAMSSWAPPLERSKLLSISYAGAQLGTVVSLPLSGVICYNNWTVVFFFGIVGLI 240
Db 181 AMHAMSSWAPPLERSKLLSISYAGAQLGTVVSLPLSGVICYNNWTVVFFFGIVGLI 240
Qy 241 FILMICLVSDTPETHKTTTPEKEKYLSSKQKSSQKSVPMPLKSLPLWAIWVAHFS 300
Db 241 FILMICLVSDTPETHKTTTPEKEKYLSSKQKSSQKSVPMPLKSLPLWAIWVAHFS 300
Qy 301 YNNWTFYLLTLLPTMYKEVLRFNIOENGFLSAVPYLGWCMLSGQAADNLRARNWFST 360
Db 301 YNNWTFYLLTLLPTMYKEVLRFNIOENGFLSAVPYLGWCMLSGQAADNLRARNWFST 360
Qy 361 LWRRVFSLIGMIGPAIFLVAAGFIGCDYSIAVAFITISTITLGGFCSSGFSINHLDIAPS 420
Db 361 LWRRVFSLIGMIGPAIFLVAAGFIGCDYSIAVAFITISTITLGGFCSSGFSINHLDIAPS 420
Qy 421 YAGILLGITNTFATIPGMIGPIIARSITPENTIGEMOTVFCIAAANVFGAIFFTLPAKG 480
Db 421 YAGILLGITNTFATIPGMIGPIIARSITPENTIGEMOTVFCIAAANVFGAIFFTLPAKG 480
Qy 481 EVQNWAISSDHQHNRN 495
Db 481 EVQNWAISSDHQHNRN 495

RESULT 2

US-10-823-506-4
; Sequence 4, Application US/10823506
; Publication No. US20050002931A1
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/0105
; CURRENT APPLICATION NUMBER: US/10/823,506
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: US/09/359,167
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60-693,843
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Ovis sp.
US-10-823-506-4

Query Match 100.0%; Score 2617; DB 17; Length 495;
Best Local Similarity 100.0%; Pred. No. 9.3e-238;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKSPVSLAPSDGEGSDRTPLQAPRAEPAPVCCSARYNLAFSLPFGFVLYSLRVNL 60
Db 1 MKSPVSLAPSDGEGSDRTPLQAPRAEPAPVCCSARYNLAFSLPFGFVLYSLRVNL 60
Qy 61 SVALVDMVDSNTTAKDNRTSYECAHSAPIKVLHNQTKKYRWDATQGWILGSFFYGI 120
Db 61 SVALVDMVDSNTTAKDNRTSYECAHSAPIKVLHNQTKKYRWDATQGWILGSFFYGI 120
Qy 121 ITQIPGGVASSRGKLLGFGIFATAITFTPLAADFGVALVALRALEGLGEGVTYP 180
Db 121 ITQIPGGVASSRGKLLGFGIFATAITFTPLAADFGVALVALRALEGLGEGVTYP 180
Qy 181 AMHAMSSWAPPLERSKLLSISYAGAQLGTVVSLPLSGVICYNNWTVVFFFGIVGLI 240
Db 181 AMHAMSSWAPPLERSKLLSISYAGAQLGTVVSLPLSGVICYNNWTVVFFFGIVGLI 240
Qy 241 FILMICLVSDTPETHKTTTPEKEKYLSSKQKSSQKSVPMPLKSLPLWAIWVAHFS 300
Db 241 FILMICLVSDTPETHKTTTPEKEKYLSSKQKSSQKSVPMPLKSLPLWAIWVAHFS 300

Qy 301 YNNWTFYLLTLLPTMYKEVLRFNIOENGFLSAVPYLGWCMLSGQAADNLRARNWFST 360
Db 301 YNNWTFYLLTLLPTMYKEVLRFNIOENGFLSAVPYLGWCMLSGQAADNLRARNWFST 360
Qy 361 LWRRVFSLIGMIGPAIFLVAAGFIGCDYSIAVAFITISTITLGGFCSSGFSINHLDIAPS 420
Db 361 LWRRVFSLIGMIGPAIFLVAAGFIGCDYSIAVAFITISTITLGGFCSSGFSINHLDIAPS 420
Qy 421 YAGILLGITNTFATIPGMIGPIIARSITPENTIGEMOTVFCIAAANVFGAIFFTLPAKG 480
Db 421 YAGILLGITNTFATIPGMIGPIIARSITPENTIGEMOTVFCIAAANVFGAIFFTLPAKG 480
Qy 481 EVQNWAISSDHQHNRN 495
Db 481 EVQNWAISSDHQHNRN 495

RESULT 3

US-09-915-181A-7
; Sequence 7, Application US/09915181A
; Patent No. US20020098473A1
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, ROBERT
; APPLICANT: BELLOCCHIO, ELIZABETH
; APPLICANT: FREMEAU, ROBERT
; APPLICANT: REIMER, RICHARD
; TITLE OF INVENTION: NOVEL GLUTAMATE TRANSPORTERS
; FILE REFERENCE: 305T-932610US
; CURRENT APPLICATION NUMBER: US/09/915,181A
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/220,556
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-915-181A-7

Query Match 89.0%; Score 2329; DB 9; Length 495;
Best Local Similarity 86.3%; Pred. No. 1.4e-210;
Matches 427; Conservative 39; Mismatches 29; Indels 0; Gaps 0;

Qy 1 MKSPVSLAPSDGEGSDRTPLQAPRAEPAPVCCSARYNLAFSLPFGFVLYSLRVNL 60
Db 1 MRSFVRDLARNDGEEESTDRTPLPGAPRAEAPVCCSARYNLAILAFGGFIVVALRVNL 60
Qy 61 SVALVDMVDSNTTAKDNRTSYECAHSAPIKVLHNQTKKYRWDATQGWILGSFFYGI 120
Db 61 SVALVDMVDSNTTAKDNRTSKACPEHSAPIKVHNNQTKKYQWDAETQGWILGSFFYGI 120
Qy 121 ITQIPGGVASSRGKLLGFGIFATAITFTPLAADFGVALVALRALEGLGEGVTYP 180
Db 121 ITQIPGGVASKIGGKMLLFGILGTAVLTFTPIAADLGVGLIPLVLALEGLGEGVTYP 180
Qy 181 AMHAMSSWAPPLERSKLLSISYAGQLGTVVSLPLSGVICYNNWTVVFFFGIVGLI 240
Db 181 AMHAMSSWAPPLERSKLLSISYAGQLGTVIPLSGIICYYNNWTVVFFFGITGIFW 240
Qy 241 FILMICLVSDTPETHKTTTPEKEKYLSSKQKSSQKSVPMPLKSLPLWAIWVAHFS 300
Db 241 FLLMIWLVSPTPKKHRIKSHYEKEYILSLRNQSSQKSVPMPLKSLPLWAIWVAHFS 300
Qy 301 YNNWTFYLLTLLPTMYKEVLRFNIOENGFLSAVPYLGWCMLSGQAADNLRARNWFST 360
Db 301 YNNWTFYLLTLLPTMYKEILRFNVQENGFLSSLPYLGSWLCMLSGQAADNLRARNWFST 360
Qy 361 LWRRVFSLIGMIGPAIFLVAAGFIGCDYSIAVAFITISTITLGGFCSSGFSINHLDIAPS 420
Db 361 LCVRRIFSLIGMIGPAVFLVAAGFIGCDYSIAVAFITISTITLGGFCSSGFSINHLDIAPS 420

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:00:41 ; Search time 28.3269 seconds
(without alignments)
1304.461 Million cell updates/sec

Title: US-09-776-865-4
Perfect score: 2617
Sequence: 1 MKSPVSLAPSDGEGSDRT.....LFAKGEVQNWAIHQHGRN 495

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.psp.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.psp.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.psp.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.psp.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.psp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.psp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2617	100.0	495	4	US-09-359-167-4
2	2329	89.0	495	4	US-09-359-167-2
3	2329	89.0	495	4	US-09-915-181A-7
4	2329	89.0	536	4	US-09-359-167-8
5	2329	89.0	585	4	US-09-949-016-7705
6	1702	65.0	495	4	US-09-359-167-10
7	1702	65.0	495	4	US-09-359-167-12
8	979	37.4	589	4	US-09-740-041-2
9	965	36.9	578	4	US-09-915-181A-4
10	965	36.9	582	4	US-09-915-181A-3
11	965	36.5	850	4	US-09-915-181A-3
12	932	35.6	567	4	US-09-949-016-11354
13	931	35.6	560	1	US-08-647-484-2
14	931	35.6	560	1	US-08-647-481-2
15	931	35.6	560	1	US-08-430-033A-2
16	931	35.6	560	5	PCT-US96-05792-2
17	927	35.4	560	2	US-08-805-118-4
18	927	35.4	560	3	US-09-391-958-4
19	927	35.4	560	4	US-09-915-181A-5
20	874.5	33.4	497	4	US-09-949-016-6616
21	819	31.3	563	4	US-09-915-181A-6
22	819	31.3	576	3	US-08-864-785-1
23	797.5	30.5	465	4	US-09-915-181A-8
24	791.5	30.2	467	2	US-08-805-118-3
25	791.5	30.2	467	3	US-09-391-958-3
26	777	29.7	480	2	US-08-724-394A-9
27	761	29.1	552	4	US-09-270-767-45540

28	672	25.7	436	4	US-09-949-016-11448	Sequence 11448, A
29	672	25.7	470	2	US-08-724-394A-10	Sequence 10, Appl
30	603.5	23.1	401	2	US-08-805-118-1	Sequence 1, Appl
31	603.5	23.1	401	3	US-09-391-958-1	Sequence 1, Appl
32	592.5	22.6	480	2	US-08-724-394A-11	Sequence 11, Appl
33	561.5	21.5	380	4	US-09-949-016-7053	Sequence 7053, Ap
34	450	17.2	460	4	US-09-489-039A-9663	Sequence 9663, Ap
35	428.5	16.4	234	4	US-09-270-767-45527	Sequence 45527, A
36	389	14.9	186	4	US-09-270-767-45384	Sequence 45384, A
37	385.5	14.7	470	4	US-09-328-352-6912	Sequence 6912, Ap
38	365.5	14.0	455	4	US-09-489-039A-9942	Sequence 9942, Ap
39	365	13.9	439	3	US-09-172-952-14	Sequence 14, Appl
40	365	13.9	439	4	US-09-922-501-12	Sequence 12, Appl
41	360	13.8	157	4	US-09-270-767-61037	Sequence 61037, A
42	353	13.5	434	4	US-09-489-039A-13633	Sequence 13633, A
43	341	13.0	163	4	US-09-270-767-32926	Sequence 32926, A
44	341	13.0	163	4	US-09-270-767-48143	Sequence 48143, A
45	326.5	12.5	459	4	US-09-489-039A-10630	Sequence 10630, A

ALIGNMENTS

RESULT 1
US-09-359-167-4
; Sequence 4, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/01US
; CURRENT APPLICATION NUMBER: US/09/359,167
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: 60-693,843
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Ovis sp.
US-09-359-167-4

Query Match	100.0%;	Score	2617;	DB	4;	Length	495;
Best Local Similarity	100.0%;	Pred. No.	5.1e-275;				
Matches	495;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MKSPVSLAPSDGEGSDRTPLQORAPRAPVCCSARVNLAFSPGFFVLSRVNL	60				
Db	1	MKSPVSLAPSDGEGSDRTPLQORAPRAPVCCSARVNLAFSPGFFVLSRVNL	60				
Qy	61	SVALVDVMDNTTAKDNRTSYCAEHSAPIKVLHNTQKGYRWDATQCGWILGSFFGYI	120				
Db	61	SVALVDVMDNTTAKDNRTSYCAEHSAPIKVLHNTQKGYRWDATQCGWILGSFFGYI	120				
Qy	121	ITQIPGGYASRSGKLLGFGIPATATFTLTPLAADFGVALVALRLEGEGVTYP	180				
Db	121	ITQIPGGYASRSGKLLGFGIPATATFTLTPLAADFGVALVALRLEGEGVTYP	180				
Qy	181	AMHAMSSWAPPLERSKLLSISVAGALQGTVSLPLSSVLCYNNWTYVFFFGIVGLIW	240				
Db	181	AMHAMSSWAPPLERSKLLSISVAGALQGTVSLPLSSVLCYNNWTYVFFFGIVGLIW	240				
Qy	241	FILWICLVSDTPTHTKTTTPEYKEYILLSSKNQSSOKSPWIPMLKSLPLWAIWAHFS	300				
Db	241	FILWICLVSDTPTHTKTTTPEYKEYILLSSKNQSSOKSPWIPMLKSLPLWAIWAHFS	300				
Qy	301	YNNFTYTLTLPTPYMKEVLRFNQENGFSAVPLGCMILSGQAADNLRARNWFS	360				
Db	301	YNNFTYTLTLPTPYMKEVLRFNQENGFSAVPLGCMILSGQAADNLRARNWFS	360				
Qy	361	LWRRVSLTGMIGTGAIPFLVNAAGFIGDYSIAVAFITISTTLGCGSCSGFSINHDIAPS	420				


```
Db 361 LWRRVFSLIGMIGPAFLVAAGFTGCDYSLAVAFLTISTTLGGFCSSGFSINHLDIAPS 420
Qy 421 YAGILLGTTNTFATIPGMIGPIIARSLLTPENTIGEWQTVFCIAAANVFGAIFTFLPAKG 480
Db 421 YAGILLGTTNTFATIPGMIGPIIARSLLTPENTIGEWQTVFCIAAANVFGAIFTFLPAKG 480
Qy 481 EVQNWAIISDHQHGRN 495
Db 481 EVQNWAIISDHQHGRN 495

RESULT 2
US-09-359-167-2
; Sequence 2, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/01US
; CURRENT APPLICATION NUMBER: US/09/359,167
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: 60-693,843
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-359-167-2

Query Match 89.0%; Score 2329; DB 4; Length 495;
Best Local Similarity 86.3%; Pred. No. 9.5e-244;
Matches 427; Conservative 39; Mismatches 29; Indels 0; Gaps 0;

Qy 1 MKSPVSLAPSDGEGSDRTPLQORAPRAEPAPVCCSARYNLAFLSFFGFVLSLAVNL 60
Db 1 MRSFVRDLARNDGEESTDRTPLPGAPRAEAPVCCSARYNLAILAFPGFIVVALRVNL 60
Qy 61 SVALVDMVDSNTTAKDNRTSYECAHSAPIKVLHNQTKKKYRWDAAETQGWILGFFFYGI 120
Db 61 SVALVDMVDSNTTLEDNRTSKACPEHSAPIKVHNQTKKKYQWDAETQGWILGFFFYGI 120
Qy 121 ITQIPGGYVASRSGKLLGFGIFATAIFTLFTPLAADFGVGVALVALRALEGLGEGVTYP 180
Db 121 ITQIPGGYVASKIGKMLGFGILGTAVLTFTPLAADLGVGPLIVLRALGLEGLGEGVTFP 180
Qy 181 AMHAMWSSWAPPLERSKLLSIYAGAQLGTIVSLPLSGVICYNNWTVYVFFFGIVGIW 240
Db 181 AMHAMWSSWAPPLERSKLLSIYAGAQLGTIVSLPLSGIICYNNWTVYVFFFGTIGIFW 240
Qy 241 FILMICLVSDTPETHKTTITPEKEYIILSSLNQSSQKVPWIPMLKSLPLWAIWVAHFS 300
Db 241 FLLMIWLVSDTPQKHRIKSHYEKEYIILSSLNQSSQKVPWIPILKSLPLWAIWVAHFS 300
Qy 301 YNMTFYTLTLLPTMYKEILRFNVOENGFLSSLPYLSGLWLCMILSGQAADNLRARNFST 360
Db 301 YNMTFYTLTLLPTMYKEILRFNVOENGFLSSLPYLSGLWLCMILSGQAADNLRARNFST 360
Qy 361 LWRRVFSLIGMIGPAFLVAAGFTGCDYSLAVAFLTISTTLGGFCSSGFSINHLDIAPS 420
Db 361 LCVRRIISLIGMIGPAFLVAAGFTGCDYSLAVAFLTISTTLGGFCSSGFSINHLDIAPS 420
Qy 421 YAGILLGTTNTFATIPGMIGPIIARSLLTPENTIGEWQTVFCIAAANVFGAIFTFLPAKG 480
Db 421 YAGILLGTTNTFATIPGMIGPIIARSLLTPENTIGEWQTVFCIAAANVFGAIFTFLPAKG 480
Qy 481 EVQNWAIISDHQHGRN 495
Db 481 EVQNWALNDHGHRRH 495

RESULT 4
US-09-359-167-8
; Sequence 8, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
```

```
RESULT 3
US-09-915-181A-7
; Sequence 7, Application US/09915181A
; Patent No. 6818391
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, ROBERT
; APPLICANT: BELLOCCHIO, ELIZABETH
; APPLICANT: FREMEAU, ROBERT
; APPLICANT: REIMER, RICHARD
; TITLE OF INVENTION: NOVEL GLUTAMATE TRANSPORTERS
; FILE REFERENCE: 305T-9326100S
; CURRENT APPLICATION NUMBER: US/09/915,181A
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/220,556
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-915-181A-7

Query Match 89.0%; Score 2329; DB 4; Length 495;
Best Local Similarity 86.3%; Pred. No. 9.5e-244;
Matches 427; Conservative 39; Mismatches 29; Indels 0; Gaps 0;

Qy 1 MKSPVSLAPSDGEGSDRTPLQORAPRAEPAPVCCSARYNLAFLSFFGFVLSLAVNL 60
Db 1 MRSFVRDLARNDGEESTDRTPLPGAPRAEAPVCCSARYNLAILAFPGFIVVALRVNL 60
Qy 61 SVALVDMVDSNTTAKDNRTSYECAHSAPIKVLHNQTKKKYRWDAAETQGWILGFFFYGI 120
Db 61 SVALVDMVDSNTTLEDNRTSKACPEHSAPIKVHNQTKKKYQWDAETQGWILGFFFYGI 120
Qy 121 ITQIPGGYVASRSGKLLGFGIFATAIFTLFTPLAADFGVGVALVALRALEGLGEGVTYP 180
Db 121 ITQIPGGYVASKIGKMLGFGILGTAVLTFTPLAADLGVGPLIVLRALGLEGLGEGVTFP 180
Qy 181 AMHAMWSSWAPPLERSKLLSIYAGAQLGTIVSLPLSGVICYNNWTVYVFFFGIVGIW 240
Db 181 AMHAMWSSWAPPLERSKLLSIYAGAQLGTIVSLPLSGIICYNNWTVYVFFFGTIGIFW 240
Qy 241 FILMICLVSDTPETHKTTITPEKEYIILSSLNQSSQKVPWIPMLKSLPLWAIWVAHFS 300
Db 241 FLLMIWLVSDTPQKHRIKSHYEKEYIILSSLNQSSQKVPWIPILKSLPLWAIWVAHFS 300
Qy 301 YNMTFYTLTLLPTMYKEILRFNVOENGFLSSLPYLSGLWLCMILSGQAADNLRARNFST 360
Db 301 YNMTFYTLTLLPTMYKEILRFNVOENGFLSSLPYLSGLWLCMILSGQAADNLRARNFST 360
Qy 361 LWRRVFSLIGMIGPAFLVAAGFTGCDYSLAVAFLTISTTLGGFCSSGFSINHLDIAPS 420
Db 361 LCVRRIISLIGMIGPAFLVAAGFTGCDYSLAVAFLTISTTLGGFCSSGFSINHLDIAPS 420
Qy 421 YAGILLGTTNTFATIPGMIGPIIARSLLTPENTIGEWQTVFCIAAANVFGAIFTFLPAKG 480
Db 421 YAGILLGTTNTFATIPGMIGPIIARSLLTPENTIGEWQTVFCIAAANVFGAIFTFLPAKG 480
Qy 481 EVQNWAIISDHQHGRN 495
Db 481 EVQNWALNDHGHRRH 495

RESULT 4
US-09-359-167-8
; Sequence 8, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 05:11:02 ; Search time 108.026 seconds
(without alignments)
1772.222 Million cell updates/sec

Title: US-09-776-865-4

Perfect score: 2617

Sequence: 1 MKSPVSLAPSDGEGSDRT.....LFAKGEVQNWAIHQHGRN 495

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A: Geneseqp16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2617	100.0	495	3	AAY45088 Sheep GBS
2	2617	100.0	495	4	AAE06519 Sheep gro
3	2329	89.0	495	3	AAY45087 Partial h
4	2329	89.0	495	4	AAE66967 Human AST
5	2329	89.0	495	8	ADJ75516 Marker ge
6	2329	89.0	495	8	ADR14587 Human NF-
7	2329	89.0	495	8	ADP25217 PRO polyp
8	2329	89.0	536	3	AAY45089 Human GBS
9	2329	89.0	536	4	AAE06518 Human gro
10	2322	88.0	495	4	AAE38959 Human pol
11	1702	65.0	495	3	AAY45090 Human/She
12	1544	59.0	314	4	AAE65238 Human sod
13	1386	53.0	284	5	ABP41345 Human ova
14	1318	50.4	272	5	ABE89975 Human pol
15	1232	47.1	277	4	AAE93914 Human pol
16	1232	47.1	277	8	ADL32036 Human pro
17	1217	46.5	309	4	AAE40745 Human pol
18	991.5	37.9	559	4	ABE58701 Drosophil
19	985.5	37.7	502	4	ABE65873 Drosophil
20	985.5	37.7	502	4	ABE60525 Drosophil
21	979	37.4	589	5	AAU99329 Human tra
22	979	37.4	589	6	AAO30994 Human tra
23	979	37.4	589	7	ADD01392 Human TCH
24	979	37.4	589	7	ADG88329 Human tra
25	979	37.4	589	8	ADR10021 Human pro

26	969.5	37.0	601	7	ADD01410	AdD01410 Mouse TCH
27	968.5	37.0	493	8	ADN23181	AdN23181 Bacterial
28	965	36.9	578	7	ADG88331	ADG88331 Rat trans
29	965	36.9	582	6	ABG74796	ABG74796 Rat DNFI
30	965	36.9	582	6	ABG74797	ABG74797 Murine DN
31	965	36.9	582	6	ABM04787	ABM04787 Rat Na-de
32	965	36.9	582	7	ADCI15494	ADCI15494 Mouse DNP
33	965	36.9	582	7	ADCI15492	ADCI15492 Rat DNFI
34	962.5	36.8	582	6	ABG74795	ABG74795 Human DNP
35	962.5	36.8	582	7	ADCI15490	ADCI15490 Human DNP
36	962.5	36.8	582	7	ADD01475	ADD01475 Human VGL
37	958	36.6	566	7	ADJ95072	ADJ95072 Novel NOV
38	956.5	36.5	582	4	AAE79273	AAE79273 Human pro
39	956.5	36.5	860	5	ABB07689	ABB07689 Rat gluta
40	946.5	36.2	516	4	ABB67013	ABB67013 Drosophil
41	941	36.0	529	4	ABB63684	ABB63684 Drosophil
42	932	35.6	560	6	ABG74791	ABG74791 Human BNP
43	932	35.6	560	7	ADCI15482	ADCI15482 Human BNP
44	932	35.6	560	7	ADD01474	ADD01474 Human VGL
45	932	35.6	567	4	AAO13870	AAO13870 Human pol

ALIGNMENTS

RESULT 1

RAY45088
ID AAY45088 standard; protein; 495 AA.

XX AAY45088;

DT 31-MAY-2000 (first entry)

XX Sheep GBS toxin receptor (SP55).

XX Sheep GBS toxin receptor; group B beta-haemolytic streptococci; SP55;
pathological vascularisation; cancer metastases; angiogenesis; sheep;
neovascularisation; reperfusion injury; scarring; keloid;
chronic inflammation; rheumatoid arthritis; psoriasis; neural injury;
endothelial cell proliferation; antibacterial; anti-cancer;
anti-angiogenic; anti-inflammatory; anti-arthritis; anti-psoriatic.

OS Ovis sp.

XX	Key	Location/Qualifiers
FT	Peptide	7..22
FT		/label= p55a
FT		/note= "Immunogenic peptide"
FT	Peptide	8..35
FT		/label= p56a
FT		/note= "Immunogenic peptide"
FT	Modified-site	11..14
FT		/label= CK2 phospho site
FT		/note= "Predicted phosphorylation site"
FT	Modified-site	17..19
FT		/label= PKC phospho site
FT		/note= "Predicted phosphorylation site"
FT	Modified-site	37..39
FT		/label= PKC phospho site
FT		/note= "Predicted phosphorylation site"
FT	Domain	42..58
FT		/note= "Predicted transmembrane domain"
FT	Modified-site	55..57
FT		/label= PKC phospho site
FT		/note= "Predicted phosphorylation site"
FT	Modified-site	59..62
FT		/note= "Asn is N-glycosylated"
FT	Peptide	71..84
FT		/label= p57a
FT		/note= "Immunogenic peptide"
FT	Modified-site	71..74
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	73..76

FT /label= CK2 phospho site
FT /note= "Predicted phosphorylation site"
FT 73. .75
FT /label= PKC phospho site
FT /note= "Predicted phosphorylation site"
FT 77. .80
FT /note= "Asn is N-glycosylated"
FT 79. .82
FT /label= CK2 phospho site
FT /note= "Predicted phosphorylation site"
FT 95. .98
FT /note= "Asn is N-glycosylated"
FT 97. .100
FT /note= "Putative amidation site"
FT 97. .99
FT /label= PKC phospho site
FT /note= "Predicted phosphorylation site"
FT 126. .131
FT /note= "Predicted myristylation site"
FT 135. .157
FT /note= "Outer boundary of transmembrane domain"
FT 137. .153
FT /note= "Inner boundary of transmembrane domain"
FT 142. .147
FT /note= "Predicted myristylation site"
FT 162. .167
FT /note= "Predicted myristylation site"
FT 172. .177
FT /note= "Predicted myristylation site"
FT 205. .210
FT /note= "Predicted myristylation site"
FT 209. .214
FT /note= "Predicted myristylation site"
FT 225. .228
FT /note= "Asn is N-glycosylated"
FT 226. .252
FT /note= "Outer boundary of transmembrane domain"
FT 232. .248
FT /note= "Inner boundary of transmembrane domain"
FT 254. .256
FT /label= PKC phospho site
FT /note= "Predicted phosphorylation site"
FT 259. .262
FT /label= CK2 phospho site
FT /note= "Predicted phosphorylation site"
FT 269. .271
FT /label= PKC phospho site
FT /note= "Predicted phosphorylation site"
FT 276. .278
FT /label= PKC phospho site
FT /note= "Predicted phosphorylation site"
FT 302. .305
FT /note= "Asn is N-glycosylated"
FT 328. .345
FT /note= "Outer boundary of transmembrane domain"
FT 328. .344
FT /note= "Inner boundary of transmembrane domain"
FT 337. .342
FT /note= "Predicted myristylation site"
FT 357. .360
FT /note= "Asn is N-glycosylated"
FT 365. .389
FT /note= "Outer boundary of transmembrane domain"
FT 369. .385
FT /note= "Inner boundary of transmembrane domain"
FT 386. .391
FT /note= "Predicted myristylation site"
FT 390. .407
FT /note= "Outer boundary of transmembrane domain"
FT 390. .406
FT /note= "Inner boundary of transmembrane domain"
FT 403. .408
FT /note= "Predicted myristylation site"

FT Modified-site 423. .428
FT /note= "Predicted myristylation site"
FT Modified-site 427. .432
FT /note= "Predicted myristylation site"
FT Modified-site 452. .455
FT /label= CK2 phospho site
FT /note= "Predicted phosphorylation site"
FT 456. .479
FT /note= "Outer boundary of transmembrane domain"
FT 458. .474
FT /note= "Inner boundary of transmembrane domain"
XX
FN WO200005375-A1.
XX
XX 03-FEB-2000.
XX
XX 22-JUL-1999; 99WO-US016676.
XX
PR 22-JUL-1998; 98US-0093843P.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Hellergqvist CG, Fu C;
XX
DR WPI; 2000-205377/18.
DR N-PSDB; AAZ50876.
XX
PT New polynucleotide encoding mammalian receptor for streptococcus toxin,
PT useful for diagnosis and treatment of, e.g. pneumonia in neonates.
XX
PS Claim 10; Page 22; 109pp; English.
XX
CC The present sequence is sheep GBS (group B beta-haemolytic streptococci)
CC toxin receptor (SP55). Sheep GBS toxin receptor is an integral protein
CC with seven transmembrane domains. Expression vectors comprising the
CC coding region can be transformed into host cells to express GBS toxin
CC receptor and its fragments. Detecting the receptor in tissues is used to
CC diagnose pathological vascularisation, e.g. for detecting cancer
CC metastases. GBS toxin receptors are useful for treating conditions
CC associated with pathological angiogenesis or neovascularisation
CC (specifically cancer, reperfusion injury, scarring during wound healing,
CC keloids, chronic inflammation (rheumatoid arthritis or psoriasis) or
CC neural injury), and to raise specific antibodies used for treating early
CC onset disease. Inhibitors of this receptor are useful for treating
CC pathological or hypoxia-induced endothelial cell proliferation and
CC migration
XX
SQ Sequence 495 AA;
Query Match 100.0%; Score 2617; DB 3; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.8e-273;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKSPVSDIAPSDGEGSDRTPLLQAPRAPAPVCCSARYNLAPLSFGFVLYSLRVNL 60
Db 1 MKSPVSDIAPSDGEGSDRTPLLQAPRAPAPVCCSARYNLAPLSFGFVLYSLRVNL 60
Qy 61 SVALVDMVDSNTTAKONRTSYECAHSAPIKVLHNTQKKYRWDATQGWILGSFFVGYI 120
Db 61 SVALVDMVDSNTTAKONRTSYECAHSAPIKVLHNTQKKYRWDATQGWILGSFFVGYI 120
Qy 121 ITQIPGGYVASRSGGKLLGFGIFATAIFLTPLAADFGVGVGALVALRALEGLGEGVTYP 180
Db 121 ITQIPGGYVASRSGGKLLGFGIFATAIFLTPLAADFGVGVGALVALRALEGLGEGVTYP 180
Qy 181 AMIAMSWSWAPPLERSKLLSISYAGAOIGTVVSLPLSGVTCYNNWTVYVFFPGIVGIW 240
Db 181 AMIAMSWSWAPPLERSKLLSISYAGAOIGTVVSLPLSGVTCYNNWTVYVFFPGIVGIW 240
Qy 241 FILWICLVSDTPETHKTTITPYEKEYIILSLKNOLSSOKSVPMPLKSLPLWAIWVAHFS 300
Db 241 FILWICLVSDTPETHKTTITPYEKEYIILSLKNOLSSOKSVPMPLKSLPLWAIWVAHFS 300

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 05:13:45 ; Search time 114.894 Seconds

(without alignments)
2388.931 Million cell updates/sec

Title: US-09-776-865-2

Perfect score: 2836

Sequence: 1 MAAGAMTPRPVQPARPGGF.....LFAKEGVQNALNDHGHGRH 536

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2621	92.4	495	1 S175_HUMAN	Q9nra2 homo sapien
2	2329	82.1	495	1 S175_SHEEP	Q9mzd1 ovine aries
3	2298	81.0	495	1 S175_MOUSE	Q9bn82 mus musculus
4	1051	37.1	544	2 Q7Q580	Q7Q580 anopheles g
5	996.5	35.1	559	2 Q9VY7G	Q9VY7G drosophila
6	992	35.0	502	2 Q9VDM0	Q9VDM0 drosophila
7	991	34.9	476	2 Q7Q579	Q7Q579 anopheles g
8	987	34.8	493	1 YLD2_CAEEL	Q03567 caenorhabdi
9	970	34.2	582	2 Q920B7	Q920B7 mus musculus
10	970	34.2	582	2 Q9J112	Q9J112 rattus norv
11	966.5	34.1	584	2 Q8AW47	Q8AW47 brachydanio
12	965.5	34.0	582	2 Q9P2U8	Q9P2U8 homo sapien
13	965	34.0	582	2 Q8BLE7	Q8BLE7 mus musculus
14	955	33.7	589	2 Q8NDX2	Q8NDX2 mus sapien
15	954.5	33.7	588	2 Q8K1Q1	Q8K1Q1 rattus norv
16	954.5	33.7	588	2 Q7TSP2	Q7TSP2 rattus norv
17	946.5	33.4	479	2 Q23514	Q23514 caenorhabdi
18	945	33.3	601	2 Q8BFU8	Q8BFU8 mus musculus
19	941	33.2	560	2 Q9P2U7	Q9P2U7 homo sapien
20	939.5	33.1	576	2 Q6INC8	Q6INC8 xenopus lae
21	936	33.0	560	2 Q62634	Q62634 rattus norv
22	932	32.9	560	2 Q6PCD0	Q6PCD0 homo sapien
23	928.5	32.7	529	1 PICO_DRONE	Q9V785 drosophila
24	917	32.3	483	1 PICO_DRON	Q61369 drosophila
25	913.5	32.2	576	2 Q715L3	Q715L3 xenopus lae
26	905	31.9	955	2 Q7Q367	Q7Q367 anopheles g
27	895	31.6	466	2 Q6GLH1	Q6GLH1 homo sapien
28	889.5	31.4	533	2 Q7Q1S5	Q7Q1S5 anopheles g
29	876	30.9	497	2 Q9Y2C5	Q9Y2C5 homo sapien
30	870.5	30.7	535	2 Q7PMK4	Q7PMK4 anopheles g
31	856	30.2	479	2 Q9W4G7	Q9W4G7 drosophila

32	856	30.2	498	2	Q8CJH9	Q8CJH9 rattus norv
33	852.5	30.1	485	2	Q7PSC6	Q7PSC6 anopheles g
34	852.5	30.1	496	2	Q9VKC9	Q9VKC9 drosophila
35	845	29.8	632	2	Q86P76	Q86P76 drosophila
36	838.5	29.6	481	2	Q9VJW8	Q9VJW8 drosophila
37	837.5	29.5	470	2	Q7YT27	Q7YT27 drosophila
38	822.5	29.0	453	2	Q9NKP8	Q9NKP8 drosophila
39	822.5	29.0	465	1	NP11_RABIT	Q28722 o renal sod
40	822	29.0	498	2	Q8VCY5	Q8VCY5 mus musculus
41	821.5	29.0	512	2	Q9VPX2	Q9VPX2 drosophila
42	819	28.9	475	2	Q6AZ69	Q6AZ69 rattus norv
43	814.5	28.7	465	2	Q91WV5	Q91WV5 mus musculus
44	814.5	28.7	467	2	Q7VQJ7	Q7VQJ7 sus scrofa
45	811	28.6	493	2	Q9VR44	Q9VR44 drosophila

ALIGNMENTS

RESULT 1
S175_HUMAN
ID S175_HUMAN STANDARD; PRT; 495 AA.
AC Q9NRA2; Q8NBR5; Q9UGH0;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Sialin (Solute carrier family 17 member 5) (Sodium/sialic acid cotransporter) (AST) (Membrane glycoprotein HP59).
GN Name=SLC17A5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, AND TISSUE SPECIFICITY.
RX PubMed=11751519;
RA Fu C., Bardhan S., Cetateanu N.D., Wamil B.D., Wang Y., Yan H.-P., Shi E., Carter C., Venkov C., Yakes F.M., Page D.L., Lloyd R.S., Mernaugh R.L., Hellerqvist C.G.;
RA "Identification of a novel membrane protein, HP59, with therapeutic potential as a target of tumor angiogenesis.";
RT Clin. Cancer Res. 7:4182-4194(2001).
RL [2]
RN SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, TISSUE SPECIFICITY, VARIANT SD CYS-39, AND VARIANTS ISSD 268-SER--ASN-272 DEL; ARG-183 AND ARG-334.
RA Verheijen F.W., Verbeek E., Aula N., Bearens C.E.M.T., Havelaar A.C., Joosse M., Peltonen L., Aula P., Gaijaard H., Van der Spek P.J., Mancini G.M.S.;
RA "A new gene, encoding an anion transporter, is mutated in sialic acid storage diseases.";
RT Nat. Genet. 23:462-465(1999).
RL [3]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Oabayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Ogura Y., Abe K., Kamiyama K., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamiyama K., Katsuta N., Sato K., Tanikawa M., Yamazaki K., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Horiuchi T., Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y., Togiya S., Komai F., Hara H., Oshima A., Sasaki N., Arita M., Imose N., Musashino K., Yuuki H., Ogihara A., Sasaki N., Aotaka S., Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,

RA Highigaki H., Watanabe T., Sudiama A., Takemoto M., Kawakami E.,
RA Yamazaki M., Watanabe K., Kamagaki A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiya M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama K., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Colon;
RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Alteschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krawczynski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP VARIANTS SD CVS-39 AND GLU-136, AND VARIANTS ISSD 268-SER--ASN-272
RP DEL; ARG-183; ARG-334 AND VAL-371.
RX PubMed=10947946;
RA Aul N., Salomaa P., Timonen R., Verheijen F., Mancini G.M.S.,
RA Maenasson J.-E., Aul P., Peltonen L.;
RT "The spectrum of SLC17A5-gene mutations resulting in free sialic acid-
RT storage diseases indicates some genotype-phenotype correlation.";
RL Am. J. Hum. Genet. 67:832-840(2000).
RN [6]
RP VARIANT SD CVS-39.
RX PubMed=12794687; DOI=10.1002/ajmg.a.10246;
RA Martin R.A., Slaughter R., Natowicz M., Pearlman K., Orvisky E.,
RA Krasnewich D., Kleta R., Huizing M., Gahl W.A.;
RT "Sialic acid storage disease of the Salla phenotype in American
RT monozygous twin female sibs.";
RL Am. J. Med. Genet. 120:23-27(2003).
CC -!- FUNCTION: Primary solute translocator for anionic substances;
CC particularly it is a free sialic acid transporter in the lysosomes
CC (Probable).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Lysosomal
CC (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9NRA2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9NRA2-2; Sequence=VSP_010482, VSP_010483;
CC Note=NO experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Found in fetal lung and small intestine, and
CC at lower level in fetal skin and muscle. In the adult, detected in
CC placenta, kidney and pancreas. Abundant in the endothelial cells
CC of tumors from ovary, colon, breast and lung, but is not detected
CC in endothelial cells from the corresponding normal tissues.
CC -!- DISEASE: Defects in SLC17A5 are the cause of Salla disease (SD)

CC [MIM:604369]; also known as Finnish type sialuria, a sialic acid
CC storage disease (SASD). SASDs are autosomal recessive
CC neurodegenerative disorders characterized by hypotonia, cerebellar
CC ataxia and mental retardation. They are caused by a defect in the
CC metabolism of sialic acid which results in increased urinary
CC excretion of unconjugated sialic acid, specifically N-
CC acetylneuraminic acid. Enlarged lysosomes are seen on electron
CC microscopic studies. Clinical symptoms of SD present usually at
CC age less than 1 year and progression is slow.
CC -!- DISEASE: Defects in SLC17A5 are the cause of infantile sialic acid
CC storage disorder (ISSD) [MIM:269920]; also known as N-
CC acetylneuraminic acid storage disease (NSD). ISSD is a severe form
CC of sialic acid storage disease. Affected newborns exhibit
CC visceromegaly, coarse features and failure to thrive immediately
CC after birth. These patients have a shortened life span, usually
CC less than 2 years.
CC -!- SIMILARITY: Belongs to the major facilitator (MFS) superfamily.
CC Sodium/anion cotransporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF244577; AA97769.1; ALT_INIT.
CC EMBL; AJ387747; CAB62540.1; -.
CC EMBL; AK075320; BAC11546.1; -.
CC EMBL; BC020961; AAH20961.1; -.
CC Genew; HGNC:10933; SLC17A5.
CC MIM; 604322; -.
CC MIM; 269920; -.
CC MIM; 604369; -.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC Pfam; PF00083; Sugar tr; 1.
CC PROSITE; PS50850; MFS; 1.
CC Alternative splicing; Disease mutation; Lysosome; Symptot;
CC Transmembrane; transport.
CC TRANSMEM 42 62 Potential.
CC TRANSMEM 110 130 Potential.
CC TRANSMEM 137 157 Potential.
CC TRANSMEM 159 179 Potential.
CC TRANSMEM 201 221 Potential.
CC TRANSMEM 228 248 Potential.
CC TRANSMEM 280 300 Potential.
CC TRANSMEM 329 349 Potential.
CC TRANSMEM 366 386 Potential.
CC TRANSMEM 392 412 Potential.
CC TRANSMEM 424 444 Potential.
CC TRANSMEM 458 478 Potential.
CC VARSPPLIC 274 276 Potential.
CC VARSPPLIC 278 495 Missing (in isoform 2).
CC FT Missing (in isoform 2).
CC FT /FTid=VSP_010483.
CC FT R -> C (in SD; frequent mutation in
CC FT Finland).
CC FT /FTid=VAR_018684.
CC FT K -> E (in SD).
CC FT /FTid=VAR_018685.
CC FT H -> R (in ISSD).
CC FT /FTid=VAR_018686.
CC FT Missing (in ISSD).
CC FT /FTid=VAR_018687.
CC FT P -> R (in ISSD).
CC FT /FTid=VAR_018688.
CC FT G -> V (in ISSD).
CC FT /FTid=VAR_018689.
CC FT /FTid=VAR_018690.
CC SQ SEQUENCE 495 AA; 54639 MW; 5C6C154B3E93A19E CRC64;
CC Query Match 92.4%; Score 2621; DB 1; Length 495;

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:16:27 ; Search time 27.0339 Seconds
(without alignments)
1907.682 Million cell updates/sec

Title: US-09-776-865-2

Perfect score: 2836

Sequence: 1 MAAGAMTPRPVQPARPGGF.....LFAKGEVQWALNDHGHHRH 536

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	987	34.8	493	2 G88553	protein C38C10.2 [
2	967	34.1	472	2 S28286	hypothetical prote
3	936	33.0	560	2 I59302	brain specific Na+
4	822.5	29.0	465	2 A56410	sodium/phosphate t
5	809.5	28.5	465	2 S69915	sodium-phosphate t
6	808	28.5	563	2 T43650	probable sodium-de
7	808	28.5	576	2 H88548	protein ZK512.6 [l
8	801.5	28.3	467	2 A48916	sodium phosphate t
9	797.5	28.1	573	2 T23589	hypothetical prote
10	786.5	27.7	465	2 I39473	Na+-dependent phos
11	767.5	27.1	544	2 T24633	hypothetical prote
12	747	26.3	512	2 H84698	hypothetical prote
13	629.5	22.2	501	2 B89135	protein F25G6.7 [l
14	562	19.8	413	2 T01534	hypothetical prote
15	531	18.7	537	2 T45634	hypothetical prote
16	530.5	18.7	466	2 S40767	hypothetical prote
17	492.5	17.4	561	2 D84800	hypothetical prote
18	486.5	17.2	568	2 S44742	C02C2.4 protein -
19	476.5	16.8	592	2 T25419	hypothetical prote
20	460.5	16.2	516	2 T24729	hypothetical prote
21	452	15.9	530	2 T29418	hypothetical prote
22	445	15.7	380	2 T29248	hypothetical prote
23	444.5	15.7	543	2 T32496	hypothetical prote
24	432.5	15.3	499	2 T15201	hypothetical prote
25	431	15.2	506	2 T29968	hypothetical prote
26	419.5	14.8	493	2 T25357	hypothetical prote
27	392.5	13.8	478	2 T33942	hypothetical prote
28	389.5	13.7	445	2 T23590	hypothetical prote
29	388.5	13.7	462	2 T34365	hypothetical prote

30	386.5	13.6	473	2 T31717	hypothetical prote
31	380.5	13.4	425	2 A90055	hypothetical prote
32	372	13.1	420	2 S44900	ZK652.10 protein -
33	370.5	13.1	428	2 AD0675	membrane transport
34	362	12.8	485	2 T24115	hypothetical prote
35	360	12.7	455	1 H69752	probable glucarat
36	357.5	12.6	455	2 T34366	hypothetical prote
37	357	12.6	452	2 AE0861	probable glucarate
38	355.5	12.5	559	2 T33557	hypothetical prote
39	353	12.4	516	2 T27092	hypothetical prote
40	349	12.3	450	2 A65061	probable glucarate
41	349	12.3	450	2 C85930	probable transport
42	349	12.3	450	2 A91085	probable transport
43	339.5	12.0	445	1 D65171	hypothetical 48.8
44	336.5	11.9	493	2 T19383	hypothetical prote
45	334	11.8	487	2 T23776	hypothetical prote

ALIGNMENTS

RESULT 1

G88553

protein C38C10.2 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: G88553

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: G88553

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-493 <STO>

A:Cross-references: UNIPROT:Q03567; GB:chr_III; PIDN:CAA79549.1; PID:G3874873; GSPDB:GNO0

C:Genetics:

A:Gene: C38C10.2

A:Map position: 3

Query Match 34.8%; Score 987; DB 2; Length 493;
Best Local Similarity 39.7%; Pred. No. 4.5e-70;
Matches 194; Conservative 94; Mismatches 179; Indels 22; Gaps 6;

QY 55 BESTDRTPLLPACPAEAAAPCCSARYNLAILAFPGFIVVALRVNLVALVDMVDSNTT 114

Db 2 EGATTKPLVP-----STRFALSLVMFFGCLVTYMTNMFSAVVCVMNENKT 49

QY 115 LEDNRTSKACPEHSAPIKVNHNQTKKYQWDAETQGTWILSGFFYGYIITQIPGGYVASKI 174

Db 50 DTGVEKVSRCGKEMTPVESNSVIG-EPDWDKQTTGMVLSSFFYGYIGSIQIIGHLASRY 108

QY 175 GKMLLGFILGTAVLTFTPIAADLGVPLIVIRALEGLGEGVTFPRAMHAMSSWAPPL 234

Db 109 GKKWFWVTILGSAALLTLNPAARTSEYALAILRAAIGFLQGAFFPAHMTWMSVWGPP 168

QY 235 ERSKLLTSYAGAOLGTIVISLPLSGIICY--NMNVTYVFFYFGTIGIFWFLWLWLVS 291

Db 169 ELSVLTGYTYAGAIQGNIVLPLSGFLCEYFGDGWPSIFYIIGVGLWTAVMWYVSSD 228

QY 292 TPQKHKRISHYKEBYILSSLRNQLSSQ----KSPVPWVPLKSLPLWAIWAFHSYNWTFY 347

Db 229 KPATHPRITPEKQYIVTAVEASMGKDTGKVPSPWKILTSAPWACWAGHAFGAGW 288

QY 348 TLLTLLPTMYKEIIRFVQENGFLSSPLYGLSWLCLMILSGQAADNLRAKNWFSLCVR 407

Db 289 TMLVSLPSFLKDVGLGNLNLGAVASIPYIAFLAINAGGVLADTLRSKGLISTLNT 348

QY 408 FSLIGMTGPAVFLVAAGFIGDYS-LAVAFUTISTTGGFCSSGFSINHLIAPSYAGIL 466

Db 349 ANLVALIGQGFILVASGYCCGQDVLVIIFITCGMAISGLQYGFVFNLEIAPFSGTV 408

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:22:07 ; Search time 115.414 Seconds
(without alignments)
1794.174 Million cell updates/sec

Title: US-09-776-865-2
Perfect score: 2836
Sequence: 1 MAAGMTTPRPVQPARPGGF.....LPAKGEVQNALNDHGHHRH 536

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1726216 seqs, 386330316 residues

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pcp.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pcp.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2836	100.0	536	9	US-09-776-865-2
2	2836	100.0	536	17	US-10-823-506-8
3	2621	92.4	495	9	US-09-915-181A-7
4	2621	92.4	495	16	US-10-755-889-588
5	2621	92.4	495	17	US-10-823-506-2
6	2329	82.1	495	9	US-09-776-865-4
7	2329	82.1	495	17	US-10-823-506-4
8	1698	59.9	495	17	US-10-823-506-10
9	1698	59.9	495	17	US-10-823-506-12
10	1521	53.6	284	15	US-10-264-049-2477
11	1451	51.2	272	15	US-10-264-237-2351

12	987	34.8	493	15	US-10-369-493-5834	Sequence 5834, Ap
13	970	34.2	578	9	US-09-740-041-4	Sequence 4, Appli
14	970	34.2	578	14	US-10-389-967-4	Sequence 4, Appli
15	970	34.2	582	9	US-09-915-181A-4	Sequence 4, Appli
16	970	34.2	582	15	US-10-205-331-7	Sequence 7, Appli
17	970	34.2	582	16	US-10-734-731-12	Sequence 12, Appl
18	970	34.2	582	16	US-10-734-731-14	Sequence 14, Appl
19	970	34.2	582	16	US-10-807-500-12	Sequence 12, Appl
20	970	34.2	582	16	US-10-807-500-14	Sequence 14, Appl
21	965.5	34.0	582	16	US-10-734-731-10	Sequence 10, Appl
22	965.5	34.0	582	16	US-10-807-500-10	Sequence 10, Appl
23	955	33.7	589	9	US-09-740-041-2	Sequence 2, Appli
24	955	33.7	589	14	US-10-389-967-2	Sequence 28, Appl
25	955	33.7	589	18	US-10-499-731-28	Sequence 46, Appl
26	945	33.3	601	18	US-10-499-731-46	Sequence 2, Appli
27	941	33.2	560	16	US-10-734-731-2	Sequence 2, Appli
28	941	33.2	560	16	US-10-807-500-2	Sequence 300, App
29	940.5	33.2	566	15	US-10-287-226-300	GENERAL INFORMA
30	940	33.1	566	15	US-10-314-790-5	Sequence 4, Appli
31	940	33.1	560	16	US-10-734-731-4	Sequence 4, Appli
32	940	33.1	560	16	US-10-807-500-4	Sequence 4, Appli
33	939	33.1	194	15	US-10-296-115-1200	Sequence 1200, Ap
34	936	33.0	560	9	US-09-991-212A-4	Sequence 4, Appli
35	936	33.0	560	9	US-09-915-181A-5	Sequence 5, Appli
36	936	33.0	560	10	US-09-965-522-4	Sequence 6, Appli
37	936	33.0	560	16	US-10-734-731-6	Sequence 6, Appli
38	936	33.0	560	16	US-10-734-731-8	Sequence 8, Appli
39	936	33.0	560	16	US-10-807-500-6	Sequence 8, Appli
40	936	33.0	560	16	US-10-807-500-8	Sequence 8, Appli
41	936	33.0	560	17	US-10-877-818-4	Sequence 4, Appli
42	935	33.0	850	9	US-09-915-181A-3	Sequence 3, Appli
43	808	28.5	563	9	US-09-915-181A-6	Sequence 6, Appli
44	808	28.5	576	14	US-10-014-079A-1	Sequence 1, Appli
45	808	28.5	576	15	US-10-369-493-5823	Sequence 5823, Ap

ALIGNMENTS

RESULT 1
US-09-776-865-2
; Sequence 2, Application US/09776865
; Patent No. US20020061846A1
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; TITLE OF INVENTION: Methods for Preventing or Attenuating Pathoangiogenic Conditions
; FILE REFERENCE: 22100-0100 46126-252687
; CURRENT APPLICATION NUMBER: US/09/776.865
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 60/179,870
; PRIOR FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-776-865-2

Query Match	100.0%	Score 2836;	DB 9;	Length 536;
Best Local Similarity	100.0%	Pred. No. 3.6e-249;		
Matches 536;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	MAAGMTTPRPVQPARPGFSGRRSLLCCVASTPAHVGMRSFVRDLARNDEESTDR	60	
Db	1	MAAGMTTPRPVQPARPGFSGRRSLLCCVASTPAHVGMRSFVRDLARNDEESTDR	60	
Oy	61	TPLLPGAPRAEAAVCCSARYNLAIFAGFPFIYALRVNLSVALVDMVDSNTTLEDR	120	
Db	61	TPLLPGAPRAEAAVCCSARYNLAIFAGFPFIYALRVNLSVALVDMVDSNTTLEDR	120	
Oy	121	SKACPEHSAPTKVHNNTGKKYQWDATQGWILGSFFGYIITOTIPGVYVASKIGKMLL	180	

Db 121 SKACPEHSAPIKVHNTQTKKYQWDAETQGWILGSFFYGYIITQIPGGYVASKIGKMLL 180
Qy 181 GFGILGTAVLTFTPIAADLGVGLIIVLRALEGLGEGVTFPAMHAMWSSWAPPLERSKLL 240
Db 181 GFGILGTAVLTFTPIAADLGVGLIIVLRALEGLGEGVTFPAMHAMWSSWAPPLERSKLL 240
Qy 241 SISVAGAOLGTVISLPLSGIICYYNNWTVYVFFGTIGIFWFLMIWLVSOTPOKHKRIS 300
Db 241 SISVAGAOLGTVISLPLSGIICYYNNWTVYVFFGTIGIFWFLMIWLVSOTPOKHKRIS 300
Qy 301 HYEKEYILSSLRNQLSSQKSPVWPVPIKLSPLWAIIVVAHFSYNNWTFYLLTLLPTMYKEI 360
Db 301 HYEKEYILSSLRNQLSSQKSPVWPVPIKLSPLWAIIVVAHFSYNNWTFYLLTLLPTMYKEI 360
Qy 361 LRFNVQENGFLSSLPYLGSWLCMILSGQAADNLRAKNFSTLCVRRIFSLIGMIGPAVFL 420
Db 361 LRFNVQENGFLSSLPYLGSWLCMILSGQAADNLRAKNFSTLCVRRIFSLIGMIGPAVFL 420
Qy 421 VAAGFIGCDYSLAVAFITITSTLGGFCSSGFSINHLDIAPSYAGILLGITNTFTATIPGMV 480
Db 421 VAAGFIGCDYSLAVAFITITSTLGGFCSSGFSINHLDIAPSYAGILLGITNTFTATIPGMV 480
Qy 481 GPVIAKSLTPDNTVGEWQTVFYIAAINVFGAIFFTLFAKGEVQNWALNDHHGHRH 536
Db 481 GPVIAKSLTPDNTVGEWQTVFYIAAINVFGAIFFTLFAKGEVQNWALNDHHGHRH 536

RESULT 2

US-10-823-506-8
; Sequence 8, Application US/10823306
; Publication No. US2005000931A1
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/0105
; CURRENT APPLICATION NUMBER: US/10/823,506
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: US/09/359,167
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60-693,843
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 8
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-823-506-8

Query Match 100.0%; Score 2836; DB 17; Length 536;
Best Local Similarity 100.0%; Pred. No. 3.6e-249;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAAGAMTPPRVQPARPGGFLSGRRSLLCQVASTPAHVGMRSVPRDLARNDGEESTDR 60
Db 1 MAAGAMTPPRVQPARPGGFLSGRRSLLCQVASTPAHVGMRSVPRDLARNDGEESTDR 60
Qy 61 TPLLPAPRAEAPVCCSARYNLAIFAFGGFIVVLRVNLVSVLVDMVDSNTTLEDNRT 120
Db 61 TPLLPAPRAEAPVCCSARYNLAIFAFGGFIVVLRVNLVSVLVDMVDSNTTLEDNRT 120
Qy 121 SKACPEHSAPIKVHNTQTKKYQWDAETQGWILGSFFYGYIITQIPGGYVASKIGKMLL 180
Db 121 SKACPEHSAPIKVHNTQTKKYQWDAETQGWILGSFFYGYIITQIPGGYVASKIGKMLL 180
Qy 181 GFGILGTAVLTFTPIAADLGVGLIIVLRALEGLGEGVTFPAMHAMWSSWAPPLERSKLL 240
Db 181 GFGILGTAVLTFTPIAADLGVGLIIVLRALEGLGEGVTFPAMHAMWSSWAPPLERSKLL 240
Qy 241 SISVAGAOLGTVISLPLSGIICYYNNWTVYVFFGTIGIFWFLMIWLVSOTPOKHKRIS 300
Db 241 SISVAGAOLGTVISLPLSGIICYYNNWTVYVFFGTIGIFWFLMIWLVSOTPOKHKRIS 300

Qy 301 HYEKEYILSSLRNQLSSQKSPVWPVPIKLSPLWAIIVVAHFSYNNWTFYLLTLLPTMYKEI 360
Db 301 HYEKEYILSSLRNQLSSQKSPVWPVPIKLSPLWAIIVVAHFSYNNWTFYLLTLLPTMYKEI 360
Qy 361 LRFNVQENGFLSSLPYLGSWLCMILSGQAADNLRAKNFSTLCVRRIFSLIGMIGPAVFL 420
Db 361 LRFNVQENGFLSSLPYLGSWLCMILSGQAADNLRAKNFSTLCVRRIFSLIGMIGPAVFL 420
Qy 421 VAAGFIGCDYSLAVAFITITSTLGGFCSSGFSINHLDIAPSYAGILLGITNTFTATIPGMV 480
Db 421 VAAGFIGCDYSLAVAFITITSTLGGFCSSGFSINHLDIAPSYAGILLGITNTFTATIPGMV 480
Qy 481 GPVIAKSLTPDNTVGEWQTVFYIAAINVFGAIFFTLFAKGEVQNWALNDHHGHRH 536
Db 481 GPVIAKSLTPDNTVGEWQTVFYIAAINVFGAIFFTLFAKGEVQNWALNDHHGHRH 536

RESULT 3
US-09-915-181A-7
; Sequence 7, Application US/09915181A
; Patent No. US20020098473A1
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, ROBERT
; APPLICANT: BELLOCCHIO, ELIZABETH
; APPLICANT: FREMEAU, ROBERT
; APPLICANT: REIMER, RICHARD
; TITLE OF INVENTION: NOVEL GLUTAMATE TRANSPORTERS
; FILE REFERENCE: 305T-932610US
; CURRENT APPLICATION NUMBER: US/09/915,181A
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/220,556
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 7
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-915-181A-7

Query Match 92.4%; Score 2621; DB 9; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.2e-229;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 42 MRSVPRDLARNDGEESTDRTPLLPGAPRAEAPVCCSARYNLAIFAFGGFIVVLRVNL 101
Db 1 MRSVPRDLARNDGEESTDRTPLLPGAPRAEAPVCCSARYNLAIFAFGGFIVVLRVNL 60
Qy 102 SVALVDMVDSNTTLEDNRTSKACPEHSAPIKVHNTQTKKYQWDAETQGWILGSFFYGYI 161
Db 61 SVALVDMVDSNTTLEDNRTSKACPEHSAPIKVHNTQTKKYQWDAETQGWILGSFFYGYI 120
Qy 162 ITQIPGGYVASKIGKMLLGFILGTAVLTFTPIAADLGVGLIIVLRALEGLGEGVTFP 221
Db 121 ITQIPGGYVASKIGKMLLGFILGTAVLTFTPIAADLGVGLIIVLRALEGLGEGVTFP 180
Qy 222 AMHAMWSSWAPPLERSKLLSISVAGAOLGTVISLPLSGIICYYNNWTVYVFFGTIGIFW 281
Db 181 AMHAMWSSWAPPLERSKLLSISVAGAOLGTVISLPLSGIICYYNNWTVYVFFGTIGIFW 240
Qy 282 FLMIWLVSOTPOKHKRISHYEKEYILSSLRNQLSSQKSPVWPVPIKLSPLWAIIVVAHFS 341
Db 241 FLMIWLVSOTPOKHKRISHYEKEYILSSLRNQLSSQKSPVWPVPIKLSPLWAIIVVAHFS 300
Qy 342 YNWTFTYLLTLLPTMYKEILRFNVQENGFLSSLPYLGSWLCMILSGQAADNLRAKNFST 401
Db 301 YNWTFTYLLTLLPTMYKEILRFNVQENGFLSSLPYLGSWLCMILSGQAADNLRAKNFST 360
Qy 402 LCVRRIFSLIGMIGPAVFLVAAGFIGCDYSLAVAFITITSTLGGFCSSGFSINHLDIAPS 461
Db 361 LCVRRIFSLIGMIGPAVFLVAAGFIGCDYSLAVAFITITSTLGGFCSSGFSINHLDIAPS 420

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 06:00:41 ; Search time 10.6731 Seconds
(without alignments)
1304.461 Million cell updates/sec

Title: US-09-776-865-2

Perfect score: 2836

Sequence: 1 MAAGAMTPPRPVQPARPGPF.....LFAKGEVQNALNDHGHGRH 536

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pgp.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pgp.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pgp.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pgp.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pgp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pgp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2836	100.0	536	4	US-09-359-167-8
2	2804.5	98.9	585	4	US-09-949-016-7705
3	2621	92.4	495	4	US-09-359-167-2
4	2621	92.4	495	4	US-09-915-181A-7
5	2329	82.1	495	4	US-09-359-167-4
6	1698	59.9	495	4	US-09-359-167-10
7	1698	59.9	495	4	US-09-359-167-12
8	970	34.2	578	4	US-09-740-041-4
9	970	34.2	582	4	US-09-915-181A-4
10	955	33.7	589	4	US-09-740-041-2
11	941	33.2	567	4	US-09-949-016-11354
12	940	33.1	560	4	US-08-647-484-2
13	940	33.1	560	1	US-08-647-481-2
14	940	33.1	560	1	US-08-430-033A-2
15	940	33.1	560	5	PCT-US96-05792-2
16	936	33.0	560	2	US-08-805-118-4
17	936	33.0	560	3	US-09-391-958-4
18	936	33.0	560	4	US-09-915-181A-5
19	935	33.0	850	4	US-09-915-181A-3
20	876	30.9	497	4	US-09-949-016-6616
21	810	28.6	552	4	US-09-270-767-45540
22	808	28.5	563	4	US-09-915-181A-6
23	808	28.5	576	3	US-08-864-785-1
24	806.5	28.4	465	4	US-09-915-181A-8
25	801.5	28.3	467	2	US-08-805-118-3
26	801.5	28.3	467	3	US-09-391-958-3
27	793	28.0	480	2	US-08-724-334A-9

28	686	24.2	436	4	US-09-949-016-11448	Sequence 11448, A
29	686	24.2	470	2	US-08-724-394A-10	Sequence 10, Appl
30	615.5	21.7	401	2	US-08-805-118-1	Sequence 1, Appl
31	615.5	21.7	401	3	US-09-391-958-1	Sequence 1, Appl
32	592.5	20.9	480	2	US-08-724-394A-11	Sequence 11, Appl
33	562	19.8	380	4	US-09-949-016-7053	Sequence 7053, Ap
34	453	16.0	460	4	US-09-489-039A-9663	Sequence 9663, Ap
35	445.5	15.7	234	4	US-09-270-767-45527	Sequence 45527, A
36	375	13.2	186	4	US-09-270-767-45384	Sequence 45384, A
37	373	13.2	434	4	US-09-489-039A-13633	Sequence 13633, A
38	370	13.0	470	4	US-09-328-352-6912	Sequence 6912, Ap
39	369	13.0	439	3	US-09-172-952-14	Sequence 14, Appl
40	369	13.0	439	4	US-09-922-501-12	Sequence 12, Appl
41	368	13.0	157	4	US-09-270-767-61037	Sequence 61037, A
42	355.5	12.5	455	4	US-09-489-039A-9942	Sequence 9942, Ap
43	346.5	12.2	459	4	US-09-489-039A-10630	Sequence 10630, A
44	332	11.7	163	4	US-09-270-767-32926	Sequence 32926, A
45	332	11.7	163	4	US-09-270-767-48143	Sequence 48143, A

ALIGNMENTS

RESULT 1
US-09-359-167-8
; Sequence 8, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellergvist, Carl
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/01US
; CURRENT APPLICATION NUMBER: US/09/359,167
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: 60-693,843
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-359-167-8

Query Match	100.0%	Score	2836	DB	4	Length	536
Best Local Similarity	100.0%	Pred. No.	7.8e-294				
Matches	536	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MAAGAMTPPRPVQPARPGFGLSGRRSLLCOVASTPAHVGMVRSFVRDLARNDCGEESTDR	60				
Db	1	MAAGAMTPPRPVQPARPGFGLSGRRSLLCOVASTPAHVGMVRSFVRDLARNDCGEESTDR	60				
Qy	61	TPLLPGAPRAEAAAPVCCSARYNLAILAFFGFFIVYALRVNLSVALVDMVDSNTTLEDNRT	120				
Db	61	TPLLPGAPRAEAAAPVCCSARYNLAILAFFGFFIVYALRVNLSVALVDMVDSNTTLEDNRT	120				
Qy	121	SKACPEHSAPIKVHNNQTKKKYQWDAETQGWILGSPFYGIITQIPGGYVASKIGGRMLL	180				
Db	121	SKACPEHSAPIKVHNNQTKKKYQWDAETQGWILGSPFYGIITQIPGGYVASKIGGRMLL	180				
Qy	181	GFGILGTAVLTFTPIAADLGVGELIVLRALEGEGVTPPAMHAMSSWAPPLERSKLL	240				
Db	181	GFGILGTAVLTFTPIAADLGVGELIVLRALEGEGVTPPAMHAMSSWAPPLERSKLL	240				
Qy	241	SISVAGAQLGTVISLPLSGIICYYNNNTYVPFFGTIGIPWFLWLVDSPTQKHKRIS	300				
Db	241	SISVAGAQLGTVISLPLSGIICYYNNNTYVPFFGTIGIPWFLWLVDSPTQKHKRIS	300				
Qy	301	HYKEYILSLRNQLSSQKSPVWPILKSLPLMAIVVAHFYSYNNWTFYTLTLLPTFMKEI	360				
Db	301	HYKEYILSLRNQLSSQKSPVWPILKSLPLMAIVVAHFYSYNNWTFYTLTLLPTFMKEI	360				
Qy	361	LRFNQENGFLSSULPYLGSLWLCMLSCMLSLRAKWNFSTLCVRRIFSLIGMIGPAVFL	420				

```

Db 361 LRFNVQENGLSSLYLGSWLCMLSGQAADNLRAKNWFSTLCVRRIFSLIGMIGPAVFL 420
QY 421 VAAGFIGDYSLAVAFLLITSTTLGGFCSSGSINHLDIAPS VAGILLGTTNTFATIPGMV 480
Db 421 VAAGFIGDYSLAVAFLLITSTTLGGFCSSGSINHLDIAPS VAGILLGTTNTFATIPGMV 480
QY 481 GPVIAKSLTPDNTVGEMQTVFYIAAANVFGAIFFTLFAKGEVQNWALNDHHGHRH 536
Db 481 GPVIAKSLTPDNTVGEMQTVFYIAAANVFGAIFFTLFAKGEVQNWALNDHHGHRH 536

RESULT 2
US-09-949-016-7705
; Sequence 7705, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7705
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7705

Query Match 98.9%; Score 2804.5; DB 4; Length 585;
Best Local Similarity 95.9%; Pred. No. 2.1e-290;
Matches 534; Conservative 1; Mismatches 1; Indels 21; Gaps 1;

QY 1 MAAGMT-----PPRPVQPARPGGFLSGRRSLLCQVASTPAHV 39
| | | | | :
| | | | | :
Db 29 MAAGASARPLGGTAGTRRGGAVSPPRPVQPARPGGFLSGRRSLLCQVASTPAHV 88
| | | | | :
QY 40 GVMRSPVRLARNDGEESTDRTPLPGAPRAEAPVCCSARNLAILAFGFFIYALRV 99
| | | | | :
Db 89 GVMRSPVRLARNDGEESTDRTPLPGAPRAEAPVCCSARNLAILAFGFFIYALRV 148
| | | | | :
QY 100 NLSVALVDMVDSNTTLEDNRTSKACPEHSAPIKVHNNOTGKKYQWDAETQGMILGSFFYG 159
| | | | | :
Db 149 NLSVALVDMVDSNTTLEDNRTSKACPEHSAPIKVHNNOTGKKYQWDAETQGMILGSFFYG 208
| | | | | :
QY 160 YIIQIPGGYVASKIGGKMLLGFILGTAVLTFTPIAADLGVGPLIVLRALEGIGEV 219
| | | | | :
Db 209 YIIQIPGGYVASKIGGKMLLGFILGTAVLTFTPIAADLGVGPLIVLRALEGIGEV 268
| | | | | :
QY 220 FPMHAMSSWAPPLERSKLLSISYAGAQGTVISLPSGLIICYNNWTVYFFFGTIGI 279
| | | | | :
Db 269 FPMHAMSSWAPPLERSKLLSISYAGAQGTVISLPSGLIICYNNWTVYFFFGTIGI 328
| | | | | :
QY 280 FFWLLIWLVS DTPQKHRI SHYEKEYILSSLRNQLSSQKSPVWPPIKSLPLWAI VVAH 339
| | | | | :
Db 329 FFWLLIWLVS DTPQKHRI SHYEKEYILSSLRNQLSSQKSPVWPPIKSLPLWAI VVAH 388
| | | | | :
QY 340 FSYNWTFTLLTLLPTMKEILRFNVQENGFLSSLPYLSWLCMLSCQAADNLRKNP 399
| | | | | :
Db 389 FSYNWTFTLLTLLPTMKEILRFNVQENGFLSSLPYLSWLCMLSCQAADNLRKNP 448
| | | | | :
QY 400 STL CVRRIFSLIGMIGPAVFLVAAGFIGDYSLAVAFLLITSTTLGGFCSSGFSINHL DIA 459
| | | | | :
Db 449 STL CVRRIFSLIGMIGPAVFLVAAGFIGDYSLAVAFLLITSTTLGGFCSSGFSINHL DIA 508
| | | | | :


```

```

QY 460 PSYAGILLGTTNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAANVFGAIFFTLFA 519
| | | | | :
Db 509 PSYAGILLGTTNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAANVFGAIFFTLFA 568
| | | | | :
QY 520 KGEVQNWALNDHHGHRH 536
| | | | | :
Db 569 KGEVQNWALNDHHGHRH 585
| | | | | :

RESULT 3
US-09-359-167-2
; Sequence 2, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Pu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/0105
; CURRENT APPLICATION NUMBER: US/09/359,167
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: 60-693,843
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 2
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-359-167-2

Query Match 92.4%; Score 2621; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 6.4e-271;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 MRSVPRDLARNDGEESTDRTPLPGAPRAEAPVCCSARNLAILAFGFFIYALRVNL 101
| | | | | :
Db 1 MRSVPRDLARNDGEESTDRTPLPGAPRAEAPVCCSARNLAILAFGFFIYALRVNL 60
| | | | | :
QY 102 SVALVDMVDSNTTLEDNRTSKACPEHSAPIKVHNNOTGKKYQWDAETQGMILGSFFYGYI 161
| | | | | :
Db 61 SVALVDMVDSNTTLEDNRTSKACPEHSAPIKVHNNOTGKKYQWDAETQGMILGSFFYGYI 120
| | | | | :
QY 162 ITQIPGGYVASKIGGKMLLGFILGTAVLTFTPIAADLGVGPLIVLRALEGIGEV 221
| | | | | :
Db 121 ITQIPGGYVASKIGGKMLLGFILGTAVLTFTPIAADLGVGPLIVLRALEGIGEV 180
| | | | | :
QY 222 AMHAMSSWAPPLERSKLLSISYAGAQGTVISLPSGLIICYNNWTVYFFFGTIGIFW 281
| | | | | :
Db 181 AMHAMSSWAPPLERSKLLSISYAGAQGTVISLPSGLIICYNNWTVYFFFGTIGIFW 240
| | | | | :
QY 282 FLLMIWLVS DTPQKHRI SHYEKEYILSSLRNQLSSQKSPVWPPIKSLPLWAI VVAHFS 341
| | | | | :
Db 241 FLLMIWLVS DTPQKHRI SHYEKEYILSSLRNQLSSQKSPVWPPIKSLPLWAI VVAHFS 300
| | | | | :
QY 342 YNWTFTYLLTLLPTMKEILRFNVQENGFLSSLPYLSWLCMLSGQAADNLRKNPST 401
| | | | | :
Db 301 YNWTFTYLLTLLPTMKEILRFNVQENGFLSSLPYLSWLCMLSGQAADNLRKNPST 360
| | | | | :
QY 402 LCVRRIFSLIGMIGPAVFLVAAGFIGDYSLAVAFLLITSTTLGGFCSSGFSINHL DIAPS 461
| | | | | :
Db 361 LCVRRIFSLIGMIGPAVFLVAAGFIGDYSLAVAFLLITSTTLGGFCSSGFSINHL DIAPS 420
| | | | | :
QY 462 YAGILLGTTNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAANVFGAIFFTLFAK 521
| | | | | :
Db 421 YAGILLGTTNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAANVFGAIFFTLFAK 480
| | | | | :
QY 522 EVQNWALNDHHGHRH 536
| | | | | :
Db 481 EVQNWALNDHHGHRH 495
| | | | | :

RESULT 4
US-09-915-181A-7

```

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2836	100.0	536	3	AAy45089	AAy45089 Human GBS
2	2836	100.0	536	3	AAE06518	AAE06518 Human gro
3	2831	92.4	495	3	AAy45087	AAy45087 Partial h
4	2621	92.4	495	4	AAE666967	AAE666967 Human AST
5	2621	92.4	495	8	ADJ75516	ADJ75516 Marker ge
6	2621	92.4	495	8	ADr14587	ADr14587 Human NF-
7	2621	92.4	495	8	ADP25217	ADP25217 PRO poly
8	2614	92.2	495	4	AAm38959	AAm38959 Human pol
9	2329	82.1	495	3	AAy45088	AAy45088 Sheep GBS
10	2329	82.1	495	4	AAE06519	AAE06519 Sheep gro
11	1698	59.9	495	3	AAy45090	AAy45090 Human/She
12	1679	59.2	314	4	AAE65238	AAE65238 Human sod
13	1614	56.9	309	4	AAW40745	AAW40745 Human pol
14	1521	53.6	284	5	ABP411345	ABP411345 Human ova
15	1451	51.2	272	5	ABSB89975	ABSB89975 Human pol
16	1450	51.1	277	4	AAm93914	AAm93914 Human pro
17	1450	51.1	277	8	ADJ32036	ADJ32036 Human pro
18	996.5	35.1	559	4	ABBS58701	ABBS58701 Drosophil
19	992	35.0	502	4	ABBS65873	ABBS65873 Drosophil
20	992	35.0	502	4	ABBS60525	ABBS60525 Drosophil
21	987	34.8	493	8	ADn23181	ADn23181 Bacterial
22	970	34.2	578	7	ADG88331	ADG88331 Rat trans
23	970	34.2	582	6	ABG74796	ABG74796 Rat DNFI
24	970	34.2	582	6	ABG74797	ABG74797 Murine DN
25	970	34.2	582	6	ABM04787	ABM04787 Rat Na-de

FT Modified-site 136. .139 /note= "Asn is N-glycosylated"

FT Modified-site 138. .141 /note= "Putative amidation site"

FT Modified-site 138. .140 /label= PKC_phospho_site

FT Modified-site 167. .172 /note= "Putative phosphorylation site"

FT Modified-site 183. .188 /note= "Putative myristylation site"

FT Modified-site 213. .218 /note= "Putative myristylation site"

FT Modified-site 246. .251 /note= "Putative myristylation site"

FT Modified-site 250. .255 /note= "Putative myristylation site"

FT Modified-site 266. .269 /note= "Asn is N-glycosylated"

FT Modified-site 297. .300 /label= CAMP_phospho_site

FT Modified-site 300. .303 /label= PK2_phospho_site

FT Modified-site 310. .312 /note= "Putative phosphorylation site"

FT Modified-site 317. .320 /label= PKC_phospho_site

FT Modified-site 343. .346 /note= "Putative phosphorylation site"

FT Modified-site 378. .383 /note= "Asn is N-glycosylated"

FT Modified-site 398. .401 /note= "Putative myristylation site"

FT Modified-site 427. .432 /note= "Asn is N-glycosylated"

FT Modified-site 444. .449 /note= "Putative myristylation site"

FT Modified-site 464. .469 /note= "Putative myristylation site"

FT Modified-site 468. .473 /note= "Putative myristylation site"

FT Modified-site 493. .496 /label= CK2_phospho_site

FT Modified-site /note= "Putative phosphorylation site"

WO200005375-A1.

03-FEB-2000.

22-JUL-1999; 99WO-US016676.

22-JUL-1998; 98US-0093843P.

(UYVA-) UNIV VANDERBILT.

Hellerqvist CG, Fu C;

WPI; 2000-205377/18.

N-PSDB; AAZ50879.

New polynucleotide encoding mammalian receptor for streptococcus toxin, useful for diagnosis and treatment of, e.g. pneumonia in neonates.

Claim 10; Page 93-95; 109pp; English.

The present sequence is partial human GBS (group B beta-haemolytic streptococci) toxin receptor (HP5). GBS toxin receptor is an integral protein with seven transmembrane domains. Expression vectors comprising the coding region can be transformed into host cells to express GBS toxin

CC receptor and its fragments. Detecting the receptor in tissues is used to diagnose pathological vascularisation, e.g. for detecting cancer metastases. GBS toxin receptors are useful for treating conditions associated with pathological angiogenesis or neovascularisation (specifically cancer, reperfusion injury, scarring during wound healing, keloids, chronic inflammation (rheumatoid arthritis or psoriasis) or neural injury), and to raise specific antibodies used for treating early onset disease. Inhibitors of this receptor are useful for treating CC pathological or hypoxia-induced endothelial cell proliferation and CC migration

XX Sequence 536 AA;

Query Match 100.0%; Score 2836; DB 3; Length 536;
Best Local Similarity 100.0%; Pred. No. 1.5e-282;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAGAMTPPRPVQPARPGFGLSGRRSLLCQVASTPAHVGMRSVPVRLDARNDGEESTDR 60
Db 1 MAAGAMTPPRPVQPARPGFGLSGRRSLLCQVASTPAHVGMRSVPVRLDARNDGEESTDR 60
Qy 61 TPLLPGAPRAEAPVCCSARYNLAILAFPGFFIVYALRVNLVALVDMVDSNTTLENDRT 120
Db 61 TPLLPGAPRAEAPVCCSARYNLAILAFPGFFIVYALRVNLVALVDMVDSNTTLENDRT 120
Qy 121 SKACPEHSAPIKVHHNQTGKKYQWDAETQGMWILGSPFYGYIITQIPGGYVASKIGKMLL 180
Db 121 SKACPEHSAPIKVHHNQTGKKYQWDAETQGMWILGSPFYGYIITQIPGGYVASKIGKMLL 180
Qy 181 GFGILGTAVLTFTPIAADLGVGLIVLRALEGLGEGVTPAMHAMSSWAPPLERSKLL 240
Db 181 GFGILGTAVLTFTPIAADLGVGLIVLRALEGLGEGVTPAMHAMSSWAPPLERSKLL 240
Qy 241 SISYAGAQLGVISLPLSGIICYYMNTYVYFPFGTIGIFWLLIWLVS DTPQKHRS 300
Db 241 SISYAGAQLGVISLPLSGIICYYMNTYVYFPFGTIGIFWLLIWLVS DTPQKHRS 300
Qy 301 HYEKEYILSSRLNQLSSQKSPVWPVPLKSLPLMAIVVAHFSYNWTFYTLTLPTMKEI 360
Db 301 HYEKEYILSSRLNQLSSQKSPVWPVPLKSLPLMAIVVAHFSYNWTFYTLTLPTMKEI 360
Qy 361 LRFNVQENGFLSSLPYLGSWLWMLSCQAADNLRKWNFTLCVRRIFSLIGMIGPAVFL 420
Db 361 LRFNVQENGFLSSLPYLGSWLWMLSCQAADNLRKWNFTLCVRRIFSLIGMIGPAVFL 420
Qy 421 VAAGFIGCDYSLAVAFITISTLLGGFCSSGFSINHLDIAPSYAGILLGITNTFATIGMV 480
Db 421 VAAGFIGCDYSLAVAFITISTLLGGFCSSGFSINHLDIAPSYAGILLGITNTFATIGMV 480
Qy 481 GPVIAKSLTPDNTVGEWQTVFYIAAAINVFGAIFFTLFAKGEVQWALNDHHGHRH 536
Db 481 GPVIAKSLTPDNTVGEWQTVFYIAAAINVFGAIFFTLFAKGEVQWALNDHHGHRH 536

RESULT 2

AAE06518

ID AAR06518 standard; protein; 536 AA.

XX AAE06518;

AC AAE06518;

XX 16-OCT-2001 (first entry)

XX 16-OCT-2001 (first entry)

XX Human group B beta-haemolytic Streptococci toxin receptor (HP59) protein.

Human; group B beta-haemolytic Streptococci toxin receptor; GBS; HP59; cytosolic; vulnery; antiatherosclerotic; osteopathic; vasotropic; prevention; attenuation; pathoangiogenic condition; cancer; scar; wound healing; gliosis; nerve injury; chronic wound; reperfusion injury; keloid; rheumatoid arthritis; atherosclerosis; osteoarthritis; psoriasis; vaccine.

XX Homo sapiens.

OS

XX